153

THE BRITISH JOURNAL OF SOCIAL AND CLINICAL PSYCHOLOGY

EDITED BY

THELMA VENESS AND JACK INGHAM

REVIEW EDITORS

MAY DAVIDSON AND ALBERT CHERNS

Press Editor: ALBERT DICKSON

Bureau of Edn! & Psyl. Research
(S. C. E. R. T.)
Date 3.3.81
Acc 5.648

BUREAU OF EDUCATIONAL AND PSYCHOLOGICAL RESEARCH
GOVT. OF W. BENGAL
LIBRARY
CALCUTTA

Volume 10, 1971

CAMBRIDGE
Published for the British Psychological Society
AT THE UNIVERSITY PRESS
1971

£6.50

PUBLISHED BY
THE SYNDICS OF THE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London N.W.1.
American Branch: 32 East 57th Street, New York, N.Y. 10022

ureau of Ednl & Psyl Research
S. C. E R. T)
e ....
....

Printed in Great Britain by William Clowes & Sons Limited
London, Colchester and Beccles

CONTENTS OF VOLUME 10

PART 1. FEBRUARY 1971

GIBSON, H. B.
The factorial structure of juvenile delinquency: a study of self-reported acts *page* 1

STACEY, B. G. and GREEN, R. T.
Working-class conservatism: a review and an empirical study 10

ARGYLE, MICHAEL and MCHENRY, ROBERT
Do spectacles really affect judgements of intelligence? 27

HAWKS, D. V. and ROBINSON, K. N.
Information processing in schizophrenia: the effect of varying the rate of presentation and introducing interference 30

RUMP, E. E. and COURT, JOHN
The Eysenck Personality Inventory and social desirability response set with student and clinical groups 42

MARCHBANKS, GABRIELLE and WILLIAMS, MOYRA
The effect of speed on comprehension in schizophrenia 55

COCHRANE, RAYMOND
High blood pressure as a psychosomatic disorder: a selective review 61

FRITH, C. D.
Smoking behaviour and its relation to the smoker's immediate experience 73

RAY, JOHN J.
'A new measure of conservatism': its limitations 79

HARTLEY, JAMES and HOLT, JANET
A note on the validity of the Wilson–Patterson measure of conservatism 81

INSEL, PAUL and WILSON, GLENN D.
Measuring social attitudes in children 84

MCKISSACK, IAN J.
Conformity in Ghana 87

MILNER, A. D., BEECH, H. R. and WALKER, V. J.
Decision processes and obsessional behaviour 88

KADRI, Z. N.
The use of the MMPI for personality study of Singapore students 90

BOOK REVIEWS 92

PART 2. JUNE 1971

WOBER, MALLORY
Adapting Dawson's Traditional versus Western Attitudes Scale and presenting some new information from Africa 101

LINDSKOLD, SVENN and TEDESCHI, JAMES T.
Self-esteem and sex as factors affecting influenceability 114

WALL, TOBY D., STEPHENSON, GEOFFREY M. and SKIDMORE, COLIN
Ego-involvement and Herzberg's two-factor theory of job satisfaction: an experimental field study 123

MARINO, CHARLES
Cross-national comparisons of Catholic–Protestant creativity differences 132

PHILIP, ALISTAIR E.
Psychometric changes associated with response to drug treatment 138

BANNISTER, D., FRANSELLA, FAY and AGNEW, JOYCE
Characteristics and validity of the Grid Test of Thought Disorder 144

SLADE, P. D.
Rate of information processing in a schizophrenic and a control group: the effect of increasing task complexity 152

TEASDALE, JOHN D., SEGRAVES, R. TAYLOR and ZACUNE, JIM
'Psychoticism' in drug-users 160

NOBLE, GRANT
Some comments on the nature of delinquents' identification with television heroes, fathers and best friends 172

ALLEN, VERNON L. and LEVINE, JOHN M.
Social support and conformity: the effect of response order and differentiation from the group 181

THOMAS, D. R., SHEA, J. D. and RIGBY, R. G.
Conservatism and response to sexual humour 185

GILES, HOWARD
Ethnocentrism and the evaluation of accented speech 187

GUPTA, B. S.
Adaptation of a Hindi version of the Junior Eysenck Personality Inventory 189

MCDONALD, KENNETH G. and PARKER, ADRIAN D.
Gibson Spiral Maze and MPI scores: a comparison using normal adolescent subjects 191

LOVE, H. G. I.
Extreme responses of psychiatric patients on the Semantic Differential: an explanation for discrepant findings 193

BOOK REVIEWS 194

PART 3. SEPTEMBER 1971

EYSENCK, H. J.
Social attitudes and social class 201

GIBSON, H. B.
The validity of the Eysenck Personality Inventory studied by a technique of peer-rating item by item, and by sociometric comparisons 213

MURRAY, C.
The effects of ordinal position on measured intelligence and peer acceptance in adolescence 221

TERRY, ROGER L.
Expectancy confirmation and affectivity 228

MASLOW, CATHA, YOSELSON, KATHRYN and LONDON, HARVEY
Persuasiveness of confidence expressed via language and body language 234

MARCHBANKS, GABRIELLE and WILLIAMS, MOYRA
Factors affecting word selection by schizophrenic patients 241

STEIN, KENNETH B., ROZYNKO, VITALI and PUGH, LAWRENCE A.
The heterogeneity of personality among alcoholics 253

ORWIN, WILLIAM
A study of the reliability implications of lie scores in the Junior Eysenck Personality Inventory 260

MCPHERSON, F. M., BUCKLEY, FELICITY and DRAFFAN, JOAN
'Psychological' constructs, thought-process disorder and flattening of affect 267

BREZNITZ, SHLOMO
A study of worrying 271

GILES, HOWARD
Patterns of evaluation to R.P., South Welsh and Somerset accented speech 280

NIAS, D. K. B., WILSON, G. D. and WOODBRIDGE, JOAN M.
Test-retest results on the Conservatism Scale completed under conditions of anonymity and identification 282

BONE, RONALD N.
Interference, extraversion and paired-associate learning 284

BRIERLEY, HARRY
A fully automated intellectual test 286

BOOK REVIEWS 289

PART 4. NOVEMBER 1971

DELIA, JESSE G., GONYEA, ANDREW H. and CROCKETT, WALTER H.
The effects of subject-generated and normative constructs upon the formation of impressions 301

BROWN, GEOFFREY and JOHNSON, SUSAN P.
The attribution of behavioural connotations to shaded and white figures by Caucasian children 306

BROWN, STEVEN R. and HENDRICK, CLYDE
Introversion, extraversion and social perception 313

VINE, IAN
Judgement of direction of gaze: an interpretation of discrepant results 320

WRIGHT, DEREK and COX, EDWIN
Changes in moral belief among sixth-form boys and girls over a seven-year period in relation to religious belief, age and sex difference 332

TEASDALE, JOHN D. and BEAUMONT, J. GRAHAM
The effect of mood on performance on the Modified New Word Learning Test (Walton-Black) 342

O'CONNOR, N. and HERMELIN, B.
Inter- and intra-modal transfer in children with modality specific and general handicaps 346

WOOD, M. M.
Symptom assessment in epidemiology: a comparison between two methods 355

IRVING, GEORGE
The performance of elderly subjects on Gooddy's Arrow Drawing Test 360

EYSENCK, SYBIL B. G. and EYSENCK, HANS J.
A comparative study of criminals and matched controls on three dimensions of personality 362

GLAISTER, BRIAN R.
An ordinate comparison method of calculating brain damage probability 367

MAYO, P. R. and BELL, J. M.
Hostility and personality in a student-teacher population 375

LEISMAN, GERALD
Cognitive interference in spastic-hemiplegic children on the Stroop Word-Color Test 379

MIKESELL, RICHARD H. and PERSENSKY, J. J.
Simulation of liberal and conservative attitudes on the Conservatism Scale 383

KOENIG, FREDRICK
Positive affective stimulus value and accuracy of role perception 385

FOULDS, G. A.
The personal illness continuum 387

GREGSON, R. A. M. and GORDON, A. V.
A comment on Foulds on continua 388

BOOK REVIEWS 389

INDEX TO AUTHORS

AGNEW, J. 144
ALLEN, V. L. 181
ARGYLE, M. 27
BANNISTER, D. 144
BEAUMONT, G. 342
BEECH, H. R. 88
BELL, J. M. 375
BONE, R. N. 284
BREZNITZ, S. 271
BRIERLEY, H. 286
BROWN, G. 306
BROWN, S. R. 313
BUCKLEY, F. 267
COCHRANE, R. 61
COURT, J. 42
COX, E. 332
CROCKETT, W. H. 301
DELIA, J. G. 301
DRAFFAN, J. 267
EYSENCK, H. J. 201, 362
EYSENCK, S. B. G. 362
FOULDS, G. A. 387
FRANSELLA, F. 144
FRITH, C. D. 73
GIBSON, H. B. 1, 213
GILES, H. 187, 280
GLAISTER, B. R. 367
GONYEA, A. H. 301
GORDON, A. V. 388
GREEN, R. T. 10
GREGSON, R. A. M. 388
GUPTA, B. S. 189
HARTLEY, J. 81
HAWKS, D. V. 30
HENDRICK, C. 313
HERMELIN, B. 346
HOLT, J. 81
INSEL, P. 84
IRVING, G. 360
JOHNSON, S. P. 306
KADRI, Z. N. 90
KOENIG, F. 385
LEISMAN, G. 379
LEVINE, J. M. 181
LINDSKOLD, S. 114
LONDON, H. 234
LOVE, H. G. I. 193
MCDONALD, K. G. 191
MCHENRY, R. 27
MCKISSACK, I. J. 87
MCPHERSON, F. M. 267
MARCHBANKS, G. 55, 241
MARINO, C. 132
MASLOW, C. 234
MAYO, P. R. 375
MIKESELL, R. H. 383
MILNER, A. D. 88
MURRAY, C. 221
NIAS, D. K. B. 282
NOBLE, G. 172
O'CONNOR, N. 346
ORWIN, W. 260
PARKER, A. D. 191
PERSENSKY, J. J. 383
PHILIP, A. E. 138
PUGH, L. A. 253
RAY, J. J. 79
RIGBY, R. G. 185
ROBINSON, K. N. 30
ROZYNKO, V. 253
RUMP, E. E. 42
SEGRAVES, R. T. 160
SHEA, J. D. 185
SKIDMORE, C. 123
SLADE, P. D. 152
STACEY, B. G. 10
STEIN, K. B. 253
STEPHENSON, G. M. 123
TEASDALE, J. D. 160, 342
TEDESCHI, J. T. 114
TERRY, R. L. 228
THOMAS, D. R. 185
VINE, I. 320
WALKER, V. J. 88
WALL, T. 123
WILLIAMS, M. 55, 241
WILSON, G. D. 84, 282
WOBER, M. 101
WOOD, M. M. 355
WOODBRIDGE, J. M. 282
WRIGHT, D. 332
YOSELSON, K. 234
ZACUNE, J. 160

INDEX OF PUBLICATIONS REVIEWED

ALLEN, C. *A Textbook of Psychosexual Disorders* (2nd ed.) 296

ARNOLD, M. B. (ed.). *The Nature of Emotion: Selected Readings* 196

BALINT, M. *The Basic Fault* 199

BECK, A. T. *Depression: Clinical, Experimental and Theoretical Aspects* 395

BEECH, H. R. *Changing Man's Behaviour* 99

BERNARD, H. W. *Readings in Adolescent Development* 389

BORGATTA, E. F. & EVANS, R. R. *Smoking, Health and Behavior* 195

BOWLBY, J. *Attachment and Loss*, vol. 1: *Attachment* 93

BRILL, L. & LIEBERMAN, L. *Authority and Addiction* 391

BULL, N. J. *Moral Education* 390

BUROS, O. K. (ed.). *Personality Tests and Reviews* 295

BURGESS, R. L. & BUSHELL, D. *Behavioural Sociology: The Experimental Analysis of Social Process* 392

CARPENTER, E. & MCLUHAN, M. *Explorations in Communication* 198

CHADWICK-JONES, J. K. *Automation and Behaviour: Social Psychological Study* 92

COLE, M. & MALTZMAN, I. (eds.). *A Handbook of Contemporary Soviet Psychology* 293

COSTELLO, C. G. (ed.). *Symptoms of Psychopathology: A Handbook* 291

EMERY, F. E. *Freedom and Justice Within Walls: The Bristol Prison Experiment* 289

EYSENCK, H. J. & EYSENCK, S. B. G. *Personality Structure and Measurement* 197

GEIWITZ, P. J. *Non-Freudian Personality Theories* 96

GEORGE, F. *Models of Thinking* 290

GITCHOFF, G. T. *Kids, Cops and Kilos: A Study of Contemporary Suburban Youth* 391

HEBRON, M. E. *Motivated Learning: A Developmental Study from Birth to Senium* 196

LINGEMAN, R. R. *Drugs from A to Z: A Dictionary* 300

LONDON, P. & ROSENHAN, D. (eds.). *Foundations of Abnormal Psychology* 97

LORENZ, K. *Studies in Animal and Human Behaviour*, vol. 1 93

LUDWIG, A. O. *et al. Psychosomatic Aspects of Gynecological Disorder: Seven Psychoanalytic Case Studies* 298

LUNDIN, R. W. *Personality: A Behavioral Analysis* 96

MANN, L. *Social Psychology* 289

MARKS, I. M. *Fears and Phobias* 292

MORRIS, P. *Put Away: A Sociological Study of Institutions for the Mentally Retarded* 194

OFFER, D. *The Psychological World of the Teenager* 197

POPE, B. & SCOTT, W. H. *Psychological Diagnosis in Clinical Practice* 99

PRICE-WILLIAMS, D. R. (ed.). *Cross-cultural Studies* 195

ROBINSON, R. J. (ed.). *Brain and Early Behaviour* 397

ROY, R. & ROY. D. *Honest Sex* 294

SHAMSIE, S. J. (ed.). *Adolescent Psychiatry* 299

SKOLNIKOFF, E. B. *Science, Technology and American Foreign Policy* 95

TANNER, J. M. & INHELDER, B. (eds.). *Discussions on Child Development*, vols. 1–4 296

TOMAN, W. *Family Constellation: Its Effect on Personality and Social Behavior* (2nd ed.) 389

WILLIAMS, M. *Brain Damage and the Mind* 97

WRAGE, K. H. *Man and Woman: The Basis of Their Relationship* 298

YATES, A. J. *Behaviour Therapy* 398

British Journal of Educational Psychology

Editor: J. D. NISBET

Vol. 41 Part 2 June 1971

J. M. Ross, London, and H. R. Simpson, Harpenden: The national survey of health and development.

F. A. Haddon, Exmouth, and H. Lytton, Alberta: Primary education and divergent thinking abilities—four years on.

B. Macdonald and J. Rudduck, East Anglia: Curriculum research and development projects: barriers to success.

H. Quigley, N.F.E.R.: Reactions of eleven nursery teachers and assistants to the Peabody language development kit.

J. D. Wilson, Edinburgh: Predicting levels of first year university performance.

J. Hartley, J. Holt, Keele, and F. W. Hogarth, Newcastle-under-Lyme: Academic motivation and programmed learning.

D. M. Jeffree, Manchester, and A. Cashdan, The Open University: Severely subnormal children and their parents: an experiment in language improvement.

D. C. Munz, Saint Louis, and P. D. Jacobs, Oklahoma: An evaluation of perceived item-difficulty sequencing in academic testing.

Research Notes by A. J. Cropley and L. Clapson, Saskatchewan; A. H. Mehryar and R. Shapurian, Pahlavi; R. J. McKechnie, Leeds.

Theses Abstracts and Book Reviews

Subscription £4 *a year or* £1.50 *per part to:*
SCOTTISH ACADEMIC PRESS LIMITED
25 Perth Street, Edinburgh EH3 5DW

The Journal of SOCIAL ISSUES

1971 VOL. XXVII NO. 2

Socialization, the Law and Society

Issue Editor: June L. Tapp

Reflections . . . . June L. Tapp

The Role of Law: The Socializer
The Moral or Educative Influence of Criminal Law . . . . . . Johannes Andenaes
The Legal Threat as an Instrument of Social Change . . . Franklin Zimring and Gordon Hawkins
The "Basic Law" or "Constitution" of a Small Group . . . . . . Walter O. Weyrauch

Legal Development: The Socialized
Developing Senses of Law and Legal Justice . . . . June L. Tapp and Lawrence Kohlberg
Legal Guarantees of Individual Freedom: A Cross-National Study of the Development
of Political Thought . . . . Judith Gallatin and Joseph Adelson
Childhood Origins of Tolerance for Dissent . . . . Gail L. Zellman and David O. Sears
Socialization of Attitudes Toward the Legal System . . . . . Judith V. Torney
The Value Gap between Police and Policed . . . Milton Rokeach, Martin G. Miller, and John Snyder

Socialized Values in the Law and Society
Legal Structures and Restoring Equity . . . . Stewart Macaulay and Elaine Walster
The Idea of Right as a Social and Legal Concept . . . . . Lawrence M. Friedman
The Institutionalization of Conflict: Jury Acquittals . . . Mortimer R. Kadish and Sanford H. Kadish
"Please, Morris, Don't Make Trouble": Two Lessons in Courtroom Etiquette . . . Harry Kalven, Jr.
Biographical Sketches

Order from: Journal of Social Issues Post Office Box 1248 Ann Arbor, Michigan 48106
1971 *Annual Subscriptions Rates* INSTITUTIONS $15.00 INDIVIDUALS $9.00
SINGLE COPIES $3.00

Published quarterly by
The Society for the Psychological Study of Social Issues
(a division of the American Psychological Association)

World of Man

The Order of Things

AN ARCHAEOLOGY OF THE HUMAN SCIENCES

MICHEL FOUCAULT

Translated from the French

An analysis of the foundations of the human sciences, showing how changes within our culture since the seventeenth century have led to the emergence of 'man' as an object of knowledge. £3·75 (75s) net

The Action-Image of Society

ON CULTURAL POLITICIZATION

ALFRED WILLENER

Translated from the French by A. M. Sheridan Smith

Taking as its starting-point the accounts of student activists and observers, this sociological analysis of the events in Paris of May 1968 traces the complex interrelations of Marxism and anarchism, politics and culture, that rocked the established order and generated a new image of society. £3·75 (75s); paperback £1.75 (35s) net

Approaches to the Study of Organizational Behaviour

OPERATIONAL RESEARCH AND THE BEHAVIOURAL SCIENCES

Edited by GORDON HEALD

These theoretical and empirical papers exemplify the multidisciplinary approach to such problems of modern management as manpower planning, organizational behaviour, management information systems, environmental studies, management education, and the contribution of social scientists to organizational design £2.50 (50s) net

Behavioural Worlds

THE STUDY OF SINGLE CASES

P. G. HERBST

An exposition of the theory and practice of four techniques for the study of single cases that makes a significant contribution to the development of the behavioural sciences. £2.50 (50s) net

Towards Community Mental Health

Edited by

JOHN D. SUTHERLAND

Elizabeth Bott, Barbara Butler, Geoffrey Gorer, Elliott Jaques, Ben Morris, Herbert Phillipson, John D. Sutherland, and D. W. Winnicott consider from their different professional viewpoints ways of helping individuals and groups to meet the stresses of a technological society. £2.40 (48s) net

Family and Social Network

ROLES, NORMS, AND EXTERNAL RELATIONSHIPS IN ORDINARY URBAN FAMILIES

ELIZABETH BOTT

Second edition with a preface by Max Gluckman

'It has become a classic and is as fresh to read now as when it was first published.'
K. McDougall, Case Conference, 1968.
£2.50 (50s); paperback £1.50 (30s) net

TAVISTOCK PUBLICATIONS

British Journal of Educational Psychology

Editor: J. D. NISBET

Vol. 40 Part 3 November 1970

E. Stones, Birmingham: Verbal labelling and concept formation in primary school children.

P. H. Taylor, Birmingham: An exploratory study of teachers' perceptions of the process of planning courses.

E. Fischbein, Ileana Pampu and I. Minzat, Bucharest: Effects of age and instruction on combinatory ability in children.

H. E. Flynn and R. G. Munro, New Zealand: Evaluation of a curriculum.

M. K. Pafford, Keele: Creative activities and 'peak' experiences.

R. I. Brown and L. Semple, Canada and N. Ireland: Effects of unfamiliarity on the overt verbalisation and perceptual motor behaviour of nursery school children.

L. Cohen, Bradford: The head teacher role: some differing emphases of male and female head teachers.

A. H. Mehryar and R. Shapurian, Iran: The application of a high-grade intelligence scale in Iran.

J. Ward, Manchester: On the concept of criterion-referenced measurement.

M. M. Tatsuoka and R. B. Cattell, Illinois: Linear equations for estimating a person's occupational adjustment, based on information on occupational profiles.

Research Notes by D. Jehu, C. J. Picton and S. Futcher, Leicester; T. W. Field, N.S.W., and M. E. Poole, Australia; E. R. Brody, U.S.A.; D. S. Finlayson, Liverpool; J. McFie and J. A. Thompson, London.

Theses Abstracts and Book Reviews

Subscription £4 a year or £1 10s. per part to:

THE MANAGER, B.J.E.P. DEPARTMENT, ASSOCIATED BOOK PUBLISHERS LTD.
North Way, Andover, Hants

PSYCHIATRY

Journal for the Study of Interpersonal Processes

The Issue of Responsibility—

Some Explorations

Vol. 34, February, 1971

Gregory Bateson. The Cybernetics of "Self." A Theory of Alcoholism

D. A. Begelman. Misnaming, Metaphors, the Medical Model, and Some Muddles

Bernard Bergen and Stanley Rosenberg. The New Neo-Freudians: Psychoanalytic Dimensions of Social Change

Peter Breggin. Psychotherapy as Applied Ethics

Wolfgang Lederer. Some Moral Dilemmas Encountered in Psychotherapy

Subscriptions: $12.50, *foreign* $13.00; *payable in advance*
Single issues: $4.00

Published in February, May, August, November by
The William Alanson White Psychiatric Foundation, Inc.
1610 New Hampshire Ave., N.W., Washington, D.C. 20009
Donald L. Burnham, M.D., Editor

PUBLICATIONS OF THE JAPANESE PSYCHOLOGICAL ASSOCIATION

The Japanese Psychological Association, which was established in 1927 and is the only representative organization of Japanese Psychologists in all fields, publishes two journals, the JAPANESE JOURNAL OF PSYCHOLOGY and the JAPANESE PSYCHOLOGICAL RESEARCH.

JAPANESE JOURNAL OF PSYCHOLOGY

A Representative Journal of Psychology in Japan, including summaries in English, German or French. Bimonthly.

Contents: Original articles, Minor studies, General review, Book review, and News and Notes.

Subscription price is $8.00 per year. Single copy price is $1.50.

JAPANESE PSYCHOLOGICAL RESEARCH

A Quarterly Journal of Psychology, written in English. Issues in May, July, September and December.

Contents: Original articles, Notes and short articles.

Subscription price is $7.00 per year. We offer a special rate of $6.00 for members of IUPS member societies. Single copy price is $2.00.

Send Subscription Orders to:

JAPANESE PSYCHOLOGICAL ASSOCIATION
37-13-802, Hongo 4 chome, Bunkyo-ku, Tokyo 113, Japan

The Journal of SOCIAL ISSUES

AUTUMN 1970 VOL. XXVI NO. 4

Selected Articles: 1970
Issue Editor: Barry E. Collins

Kurt Lewin Memorial Address: 1970 . . . . . . Morton Deutsch
The Kurt Lewin Memorial Presentation . . . . . . Nevitt Sanford
Whatever Happened to Action Research
School Desegregation and Racial Cleavage: 1954–1970: . . . . Martha W. Carithers
A Review of the Literature . . . . John M. Goering and Marvin Cummins
Intervention Research and the Survey Process . . . . . Carol H. Weiss
The Politicization of Evaluation Research . . . . Jeffery M. Paige
Changing Patterns of Anti-White Attitudes Among Blacks
Three Studies of the Preferences of Students of Different Races for James Hoetker and Gary Siegel
Actors in Interracial Theatre Productions . . . . Dominic Amante,
The Epidemiological Distribution of CNS Dysfunction . . . Lewis John Weeber
Phillip H. Margules, Donna M. Hartman, Delores B. Storey, and
The Effects of an Interracial Preschool Program upon Racial Preference, Roland C. Crooks
Knowledge of Racial Differences, and Racial Identification . .
Biographical Sketches
Abstracts . . . . Nevitt Sanford
The Activists' Corner . . . . Lewis W. Brandt
Comments and Rejoinders . .

Order from: Journal of Social Issues Post Office Box 1248 Ann Arbor, Michigan 48106
1971 *Annual Subscriptions Rates* INSTITUTIONS $15.00 INDIVIDUALS $9.00
SINGLE COPIES $2.25

Published quarterly by
The Society for the Psychological Study of Social Issues
(a division of the American Psychological Association)

Br. J. soc. clin. Psychol. (1971), **10**, pp. 1–9
With 1 text-figure
Printed in Great Britain

The Factorial Structure of Juvenile Delinquency: A Study of Self-Reported Acts

By H. B. GIBSON*

Institute of Criminology, University of Cambridge

An individually administered test of self-reported delinquency was given to 402 boys aged 14–15 years, following pilot work published earlier. The results were analysed factorially and related to a measure of social handicap on which the boys has been rated some years previously, and to the boys' records of criminal conviction. A large general factor of delinquency was found and several specific factors. The general factor was related to social handicap and criminal conviction, but the specific factors showed differences in their relations to these two social variables. Second-order factors were also extracted and the first two generated a hypothesis concerning the interaction of handicap and conviction which was tested empirically on the data.

In a previous communication (Gibson, 1967) details were given of pilot work in constructing a test of self-reported delinquency by schoolboys. This followed work published in the U.S.A., Scandinavia and Britain. Mention was made of an anomaly existing with regard to the relationship between social class and self-reported delinquency. In studies where the information was gathered by an anonymous questionnaire technique the relationship appeared to be insignificant, but where the technique was that of individual interviews, the expected relationship was found—the lower the social class the greater the amount of reported delinquency. In discussing the probable reasons for this anomaly (Gibson, 1968) it was pointed out that the anonymous questionnaire technique is a somewhat unreliable means of establishing hard data about social class as well as delinquency, and a number of studies have shown that a sizable (and probably significant) minority of questionnaires have to be scrapped because of inadequate completion.

A recent British study, however (McDonald, 1969), showed quite clearly higher rates of self-reported delinquency for working-class than for middle-class boys, even though the anonymous questionnaire technique was used. Delinquency correlated more strongly with social class than with any other variable in the study. The question arises: what do we mean by 'delinquency'? It may be that there are a number of behavioural syndromes of acts, all against the law, but which have very different social class concomitants. McDonald (1969) did in fact divide the items on her inventory into a number of types of delinquency and found that relations with social class were different for the different groups of offences. This typing of offences was done entirely on common-sense face inspection of the items and not in terms of what they meant to the boys; no form of factorial analysis of the results was attempted. As what is being measured in such a test is the boys' own report of their behaviour, a legal or other classification of the acts described in the questionnaire may not be the best one in terms of how they relate to one another, and to other variables such as social class.

---

* Present address: Department of Psychological and Social Studies, The Polytechnic, Hatfield, Herts.

Gould (1969) has suggested that delinquency should be considered in terms of three separate concepts: 'delinquent acts', 'official delinquency' (i.e. those acts which have precipitated official response) and 'self-perceived delinquency'. It is the third type which Gould believes to be 'close to what the self-report scales are probably measuring; a belief that one is himself a delinquent, or at least that he has committed delinquent acts'. This is a somewhat surprising conclusion; after all, what the boy reports is whether he has committed certain acts or not, irrespective of whether he considers them to be 'delinquent' in his own value system. It may be that two acts are considered equally delinquent by adults and law-makers, but are viewed entirely differently by boys, and this difference of perception may be pronounced among boys of lower social class. One of the two acts may be freely engaged in and regarded as blameless, while the other is regarded as definitely shameful, and its perpetration may have entirely different social implications.

One of the difficulties of research into self-reported delinquency is in validating the results by external criteria. It is difficult to know how closely the boys' report of their actions accords with reality. Reiss & Rhodes (1961) interviewed a 2 per cent sample of 9000 juveniles whose court records were known; Gold (1966) tried a system of check by informants; Clark & Tifft (1966) gave follow-up interviews in which a polygraph lie-detector was employed; Gibson *et al.* (1970) examined the actual offences of which boys were convicted in relation to the different types of acts to which they admitted. Such efforts have some bearing on validity but they do not overcome the objection that people may be more able to remember and ready to admit to those of their acts which have received official or public recognition, even though they are not necessarily aware that the interviewer has knowledge of them. Where the inquiry is by anonymous questionnaires, the investigator has to fall back on the comparison of the conviction rates, etc. of criterion groups. The fact of being convicted a number of times, only once or not at all, may have important associations, both with the type of delinquent activities admitted to and with the social class of the respondent. It is therefore worth while to include it as a relevant variable in such research, even though it may not be a wholly dependable measure of the validity of self-report.

THE PRESENT STUDY

The subjects in this study were 402 boys who had been followed up since the age of 8–10 years as part of a research known as the Cambridge Study in Delinquent Development (West, 1969). At the age of 14–15 years they were interviewed individually in their schools by a psychologist, who gave them a large battery of tests. The penultimate test consisted of card-sorting and has been described in detail previously (Gibson, 1967). In the final part of this test the subject sorts 38 cards, each bearing a description of a criminal act or an act which is generally considered to be fringe-delinquent, like truanting from school or going to X films under age. To the original list of 36 acts (given in the Appendix to the 1967 paper) two items were added: 'Obtaining money by false pretences' and 'Taking illegal drugs like purple hearts or smoking reefers'. Apart from these additions to the cards, no further modifications were made to the procedure.

In addition to the data gathered about self-reported delinquency, two other pieces of information about the boys were taken into consideration:

Social handicap (*SH*). Each boy was rated on a three-point scale of SH as to whether it was 'absent' ($n = 182$), 'moderate' ($n = 168$) or 'severe' ($n = 52$). This rating had been made about four years previously; details of the scale are published elsewhere (West, 1969). The SH scale includes the question of the father's socio-economic status, and other family variables, and has been found in a number of analyses to be a good discriminative measure of grades of social class in this population which is predominantly working class.

Criminal conviction (*CC*). The boys' names were checked with the Criminal Records Office and the Children's Department for records of conviction for indictable offences. This study is based upon their complete records up to the 14th birthday (30 boys convicted) and such records as had become available after that birthday (27 additional boys convicted). CC was scored on a three-point scale: $0 =$ no conviction ($n = 345$); $1 =$ convicted once only ($n = 34$); $2 =$ convicted more than once ($n = 23$). Such records are obviously incomplete, since the delinquency status of some boys will have increased later, but less than a quarter of the sample were under 15, and nearly a quarter over 16 at the time the delinquency records were established, so most boys were past the peak age for delinquency (see McKissack, 1967). It should be noted, however, that in some cases the act for which he was convicted did not take place until after the boy had done the self-reported delinquency test. CC therefore has a general rather than a precise relationship to the self-reported record.

STATISTICAL ANALYSIS

Each act was scored 1 or 0 according to whether the boy did or did not admit to having committed it. The last item, using drugs, was dropped from the analysis, as only two boys in the population admitted doing it. The other 37 acts were intercorrelated together with the two variables SH (38) and CC (39). The resulting 39×39 matrix was factor analysed using the principal components method.

Table 1 shows the list of acts (see the 1967 study for the full text of each card) together with the frequency of denial. The table also shows the correlation of each item with the variables SH and CC, and the loading on the first unrotated eigenvector. The first unrotated eigenvector accounted for 20 per cent of the total variance—more than three times that of any single succeeding eigenvector.

Fig. 1 shows the loadings of 31 items plotted against the axes of the unrotated eigenvectors 2 and 3 which accounted for 7 per cent and 5 per cent of the total variance respectively. The eight remaining items which had greater loadings on eigenvector 4 than on either of the two preceding ones are shown in Table 2, together with the loadings of SH and CC on that component.

The first 12 principal components, accounting for 60·3 per cent of the total variance, had latent roots greater than 1, and so, following Guttman's criterion (Guttman, 1954), only these were rotated to oblique simple structure using the Promax criterion (Hendrickson & White, 1964). The analysis was continued with the extraction of second-order factors by the method described by White *et al.* (1969). Three second-order factors were extracted, which intercorrelated as follows: $r_{\mathrm{I.II}} = -0{\cdot}39$, $r_{\mathrm{I.III}} = -0{\cdot}42$, $r_{\mathrm{II.III}} = 0{\cdot}49$. The loadings of all the items on these three factors are shown in Table 3, where they are rearranged in order to indicate the factor on which each item has its principal loading.

Two important features are apparent about these second-order factors. The two key variables, which did not result from the boys' self-report, SH (38) and CC (39), may be seen to be unique among other items with regard to their loadings on I and II. The highest *positive* loading on I is SH, and the highest on II is CC.

Table 1. *List of items with their frequency, correlation with social handicap and criminal conviction, and loading on component* 1

(Decimal points omitted.)

Serial no.	Abbreviated title	Denial %	Social handicap	Conviction	Loading on 1
1	Riding cycle without lights	22·9	07	03	23
2	Driving car/motor cycle under 16	80·1	06	23	49
3	Group hooliganism	83·6	15	14	53
4	Truancy from school	45·8	15	21	44
5	Cheating on transport	28·6	−02	01	36
6	Fireworks in street	15·6	01	05	28
7	Pilfering at home	90·8	09	03	19
8	T.D.A. car/motor cycle	92·8	10	33	58
9	Vandalism—public property	88·1	12	04	48
10	Assaulting strangers	77·6	09	13	55
11	Breaking into a big store, etc.	96·3	03	30	40
12	Breaking into a small shop	93·8	10	30	45
13	Stealing out of cars	91·6	18	38	58
14	Carrying a weapon	79·4	10	15	55
15	Attacking an enemy	81·1	06	11	53
16	Breaking windows of empty houses	31.3	11	08	39
17	Using a weapon	87·1	10	24	58
18	Drinking alcohol in pubs	73·9	−11	10	38
19	Entering pubs	49·0	−03	02	38
20	Shoplifting—big stores	81·4	10	22	49
21	Shoplifting—small shops	63·4	08	23	59
22	Littering the street	75·6	10	03	41
23	Receiving stolen property	64·2	−06	10	39
24	Housebreaking—planned	95·5	22	49	55
25	Housebreaking—impulsive	94·5	24	62	57
26	Stealing a pedal cycle	91·8	20	39	57
27	Struggling to escape from police	93·3	14	28	56
28	Attacking police to free someone	98·2	12	17	33
29	Pilfering school property	70·9	−03	00	34
30	Fiddling employer's property	97·8	07	17	28
31	Trespassing	36·8	06	09	33
32	X films under age	36·1	00	10	44
33	Gambling £1 or more a week	92·3	16	14	25
34	Regularly smoking under 15	68·4	16	23	47
35	Stealing from slot machines	85·8	16	23	52
36	Stealing from people's clothes	96·7	06	12	33
37	Money by false pretences	90·1	17	13	47
38	Social handicap	—	—	24	25
39	Criminal conviction	—	24	—	48

This has emerged from a 'blind' analysis, i.e. the program was not arranged to maximize the loadings on any particular items.

The loadings of the items on the first two factors have important implications for their meaningful interpretation in the context of this study. If we compare item 20 (Shoplifting—big stores) with 34 (Smoking) they are seen to correlate 0·22, 0·23 with CC and 0·08, 0·16, with SH (Table 1). It might be supposed that both shoplifting and smoking are about equally predictive of criminal conviction in this population. However, these two items have very different loadings on I and II; item 20 loads very highly (−0·69) on I but modestly (0·24) on II, whereas item 34 loads −0·02 and 0·57 respectively on these two factors. This implies that the predictive power of item 20 for conviction is conditional on the individual boy's level of social handicap, but that item 34 is predictive irrespective of the level of SH. This hypothesis leads to an empirical test.

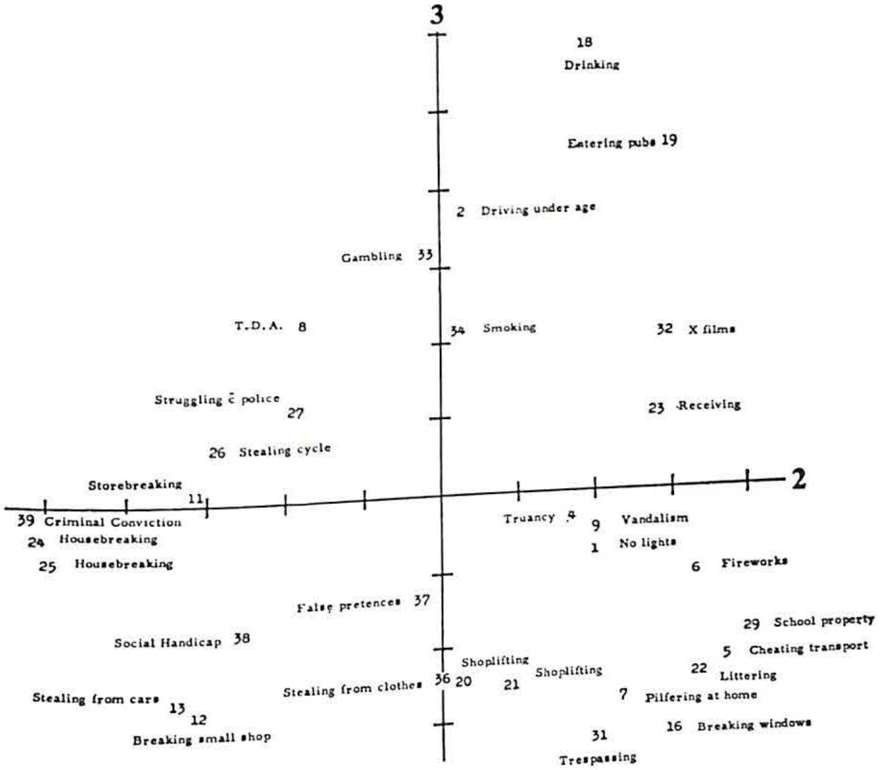


Fig. 1. Plot of the loadings of 31 items on the axes of components 2 and 3.

Table 2. *Items omitted from Fig. 1 plus SH and CC*

(Decimal points omitted.)

		Loading on 4
3	Group hooliganism	−34
30	Fiddling employer's property	−34
28	Attacking police to free someone	−34
14	Carrying a weapon	−30
15	Attacking an enemy	−28
17	Using a weapon	−21
35	Stealing from slot machines	−19
10	Assaulting strangers	−16
38	Social handicap	−22
39	Criminal conviction	08

Tables 4 and 5 show the association between CC and both items 20 and 34 respectively, but calculated separately for each of three levels of SH. It may be seen that shoplifters are convicted significantly more frequently than non-shoplifters *only* when there is a moderate or severe degree of SH. By contrast, smokers are convicted significantly more frequently when SH is absent, or present only to a moderate degree. When the degree of handicap is severe, the association is not quite significant.

Table 3. *Loadings of all items on the second-order factors*

(Decimal points omitted.)

Item	I	II	III	Item	I	II	III
21	−69	23	−13	24	−5	43	14
20	−62	24	−15	11	−6	41	−6
5	−62	−34	27	14	−16	41	9
29	−52	−30	36	35	−15	39	15
6	−46	2	−15	12	−17	37	2
32	−37	8	10	27	−5	34	29
16	−36	7	5	4	−24	32	4
31	−30	12	−1	28	−2	27	11
23	−26	12	11	30	−16	27	11
19	−21	12	13				
7	−20	1	13				
38	26	34	19	37	−7	−9	69
				9	−6	5	56
39	3	68	−15	3	13	25	48
26	−4	66	−10	22	−13	−2	45
34	−2	57	−7	33	14	−3	43
17	−6	57	5	15	−11	18	36
13	−16	54	1	36	−13	4	32
8	23	53	34	10	−14	28	27
2	20	50	24	18	−5	17	23
25	−10	49	6	1	−24	23	−28

Table 4. *Association between shoplifting (item 20) and criminal conviction at different levels of social handicap*

Social handicap	Shoplifting	Number convicted		Convicted %	χ^2 (1 d. f.)	*P*
		No	Yes			
Severe	Yes	7	11	61·1		
	No	29	6	17·1	8·63	$<0{\cdot}005$
Moderate	Yes	17	9	34·6		
	No	126	15	10·6	8·39	$<0{\cdot}005$
None	Yes	29	2	6·5		
	No	137	14	9·3	—	n.s.

Table 5. *Association between smoking (item 34) and criminal conviction at different levels of social handicap*

Social handicap	Smoking	Number convicted		Convicted %	χ^2 (1 d.f.)	P
		No	Yes			
Severe	Yes	14	12	46·2	3·46	n.s.
	No	22	5	18·5		
Moderate	Yes	42	13	23·6	4·65	$<0·05$
	No	101	11	9·8		
None	Yes	37	9	19·6	7·20	$<0·01$
	No	129	7	5·1		

DISCUSSION

Table 1 shows considerable variation in the frequency of admission to the various acts listed, as was intended in their selection. In general, the more serious offences are denied by greater percentages of boys, and have higher loadings on the first component (an unrotated eigenvector). It is suggested that component 1, accounting for 20 per cent of the total variance, is a 'general' factor of delinquency, like the factor G which emerges when a mixed battery of cognitive tests is analysed factorially. As both SH and CC have positive loadings on this 'general' factor, and are themselves positively related ($r=0·24$), it is true to say that the more handicapped boys admit to more delinquency in general, and they are more liable to be convicted; also, the boys who admit to more delinquent activity are more likely to be in the convicted category.

When we come to consider the minor components (unrotated eigenvectors) which are small in relation to the first, it must be remembered that a partialling out operation for general delinquency has taken place. Reference to Fig. 1 shows the main factorial relationships. Component 2 may be considered in terms of active theft *v.* cheating and pilfering incidental to normal activities, and disorderly acts which are not dishonest. Boys admitting to items in the 'active theft' cluster are specially liable to conviction and come from more handicapped homes. The type of acts in this negatively loading group is similar to that chosen by McDonald (1969, Table 3.1), but it is of interest to note that in the present study shoplifting did not become associated with this cluster.

Component 3, also shown on the plot of Fig. 1, can best be considered in terms of maturity *v.* immaturity. The acts loading positively on it—frequenting pubs, driving cars, gambling, etc.—are the acts which are considered delinquent in schoolboys because of their precocity. The acts loading negatively on this component are the sort of mischief which boys get up to at an early age.

Table 2 shows the acts which have higher loadings on component 4 than on 2 or 3. It will be observed that the great majority of them are acts of violence; the two property offences (30 and 35) have higher loadings on small subsidiary components. Although SH has a loading on this component in the expected direction, the loading of CC is near zero. In view of the fact that the effect of the large 'general' factor has been partialled out, it is suggested that, *holding the level of delinquency constant*, 'violence' is not specially likely to be a concomitant of conviction. It is possible that items in this category have attracted admissions out of bravado: we know of no case of a boy charged with 'attacking or fighting a policeman who is trying to arrest someone else' (item 28), although seven boys claimed to have done so.

It is natural to seek for a tidy classification of delinquent acts arising from such an analysis, but what has emerged is rather more interesting. There were as many as 12 latent roots greater than 1, and hence 12 primary factors in the oblique simple structure following rotation. The data have not been discussed in terms of these rotated primary factors because of the complexity involved; they were intercorrelated with r's ranging from 0·50 to −0·45. The number of these latent roots indicates that we cannot summarize such behavioural data easily in terms of a few orthogonal factors, although some attempt has been made to portray the more obvious features in relation to four of the eigenvectors.

The extraction of second-order factors as shown in Table 3 yielded a valuable hypothesis, and this was tested on the items concerning shoplifting and smoking. It may seem a criminological truism to demonstrate, as in Table 4, that juvenile shoplifting leads to conviction (for that or other offences) only in the more socially handicapped section of the population, but as shown in Table 5 juvenile smoking has precisely the opposite implications—it is more strongly associated with conviction among boys from better class homes. The same type of interaction between SH and CC can be predicted for the other items which have high loadings on either I or II.

Factor III is more difficult to interpret. Most of the items with high loadings on it are those of a disorderly but not necessarily dishonest character, which do not have a specially strong relationship with either SH or CC.

This factorial approach adds new insights to the study of self-reported delinquency, and departs from the earlier method of classifying acts by legal criteria or how such acts appear in the adult view. To compare such artifacts as a 'total theft score' with other variables, as done by Belson (1968), is to ignore the very real differences in psychological and sociological meaning of different types of behaviour all of which we call 'theft'. A 'total theft score' takes cognizance of a general factor but neglects the specific factors.

Some of the acts on the list are not actually illegal, and the usefulness of this approach has been questioned by McClintock (1968) who thinks that in the study of self-reported delinquency 'check-lists of crimes need to be related, at least initially, to offences as classified in the criminal statistics'. However, many fringe-delinquent acts such as trespassing are meaningful events in the lives of boys and study of their relations with criminal acts, conviction and social class may lead to greater understanding of how behaviour may best be classified meaningfully. Christie (1968) commented that 'The major problem has *not* to do with getting the informants to give honest answers, but to know what to ask them to be honest about. Basically it is a problem of sampling . . . of *acts*.' The study reported has used the original list of acts of Willcock (1965) with very little addition. Factorial study of the results may suggest modifications to such lists according to the population to be tested. Only when we know a great deal more about the various syndromes of delinquent activity will we be able to explore more effectively their relations with such variables as social class and criminal conviction.

REFERENCES

Belson, W. A. (1968). The extent of stealing by London boys and some of its origins. *Adv. Sci.* **25**, 171–184.

Christie, N. (1968). Hidden delinquency: some Scandinavian experiences. *Third National Conference on Research and Teaching in Criminology*. Cambridge: Institute of Criminology.

Gibson, H. B. (1967). Self-reported delinquency among schoolboys, and their attitudes to the police. *Br. J. soc. clin. Psychol.* **6**, 168–173.

Gibson, H. B. (1968). Self-reported delinquency: preliminary results of an ongoing study. *Third National Conference on Research and Teaching in Criminology*. Cambridge: Institute of Criminology.

Gibson, H. B., Morrison, S. & West, D. J. (1970). The confession of known offences in response to a self-reported delinquency schedule. *Br. J. Criminol.* **10**, 277–280.

Gold, M. (1966). Undetected delinquent behaviour. *J. Res. Crime Delinq.* **3**, 27–46.

Gould, L. C. (1969). Who defines delinquency: a comparison of self-reported and officially reported indices of delinquency for three racial groups. *Social Problems* **16**, 325–336.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika* **19**, 149–161.

Hendrickson, A. E. & White, P. O. (1964). Promax: a quick method for rotation to oblique simple structure. *Br. J. statist. Psychol.* **17**, 65–70.

McClintock, F. H. (1968). Criminological and penological aspects of the dark figure of crime and criminality. *Sixth European Conference of Directors of Criminological Research Institutes*. Strasbourg: Council of Europe.

McDonald, L. (1969). *Social Class and Delinquency*. London: Faber & Faber.

McKissack, I. J. (1967). The peak age for property crimes. *Br. J. Criminol.* **7**, 184–194.

Reiss, H. & Rhodes, A. L. (1961). The distribution of juvenile delinquency in the social class structure. *Am. sociol. Rev.* **26**, 720–732.

West, D. J. (1969). *Present Conduct and Future Delinquency*. London: Heinemann.

White, P. O., Soueif, M. I. & Eysenck, H. J. (1969). Factors in the Eysenck Personality Inventory. In H. J. Eysenck & S. B. G. Eysenck (eds.), *Personality Structure and Measurement*. London: Routledge & Kegan Paul.

Willcock, H. D. (1965). Deterrents and incentives to crime amongst youths aged 15 to 21 years. (Unpublished report prepared for the Home Office by the Social Survey.)

Manuscript received 19 September 1969

Revised manuscript received 18 June 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 10–26
Printed in Great Britain

Working-Class Conservatism: A Review and an Empirical Study

By B. G. STACEY

University of Strathclyde, Glasgow

and R. T. GREEN

The Open University, Bletchley

The characteristics of manual working-class Conservatives are examined. Inter-generation downward mobility is shown to make a substantial contribution to working-class Conservatism. The embourgeoisement, social pressure and deference explanations of working-class Conservatism are critically reviewed. A survey is reported which was carried out to explore the organization of attitudes in the working class. A factor analysis of 69 variables constituting a questionnaire completed by 302 subjects produced differences between working-class Conservative and Labour supporters which correspond fairly well with the patterning of attitudes in a complementary survey of white-collar subjects. The deference explanation is re-examined on the basis of these results. Lipset's analysis of working-class authoritarianism is also commented upon.

Though it is a common generalization that class is the most important single determinant of political values, attitudes and behaviour in Britain, one weakness of this generalization is plain enough. The Conservative Party receives approximately half its electoral support from the manual working class, despite the Party's overwhelmingly middle and upper class composition and its close links with the aristocracy and the 'City' (Alford, 1964; McKenzie & Silver, 1968; Runciman, 1966). The manual working class takes in two-thirds of the population and between 30 and 35 per cent of the voters in this class usually support the Conservatives.

Over the past few years several researchers have attempted to account for this major division of political loyalties in the working class and to explain working-class conservatism or traditionalism. In particular, they have attempted to elucidate the characteristics of working-class Conservatives and to distinguish them psychologically from their Labour counterparts. They have also probed the causes of the long-term electoral appeal of the Conservative Party to a sizable section of the working class. One sociologist has gone so far as to attribute the stability of the British political system to the attitudes of the working class, and particularly the Conservative sector of the working class, towards the class structure, hierarchical authority relations and the nation's governmental tradition (Nordlinger, 1967).

From various studies of the social characteristics of Conservative, Labour and Liberal supporters within the manual working class it is clear that these groups are fairly similar with regard to variables such as age, sex, income (but not personal savings), occupational skill, ownership of household goods, and church affiliation (Butler & Stokes, 1969; Goldthorpe *et al.*, 1968; Kelvin, 1965; McKenzie & Silver, 1968; Nordlinger, 1967; Runciman, 1966). Given this, however, there are certain tendencies which differentiate the Conservatives. Working-class people who consider themselves middle class are more likely to support the Conservative Party than those who see themselves as working class. Working-class people who stay on at school beyond the minimum leaving age are more likely both to describe themselves as middle class and to support the Conservative Party. Older working-class people, and particularly older women, are somewhat more likely than younger ones to vote Conservative. Among the employed working class there exists a positive association between smallness of factory or firm and the extent of right-wing attitudes and Conservative voting (Ingham, 1969). Other categories of people more likely than their counterparts to be Conservative include women, church attenders, non-union members, those brought up in small families (0–2 siblings), and those with a white-collar spouse. In general, right-wing values and attitudes are positively associated with socially heterogeneous towns, villages and communities, and negatively associated with homogeneous working-class communities.

Stacey (1966) suggested that those downwardly mobile between generations into the manual strata make a substantial contribution to the working-class Conservative vote, despite their tendency to become less Conservative. Earlier Maccoby (1954) had observed, in a sample of young Americans, that the downwardly mobile into the manual strata take on some of the labour attitudes of their stratum of arrival but remain more Republican in their party preference than their non-mobile peers. Runciman (1966) found in his national sample (England and Wales) that a good deal of the self-assessment of people in the manual strata as 'middle class' can be explained by inter-generation downward mobility, and that twice the proportion of manual strata people with fathers in white-collar jobs support the Conservative Party as manual strata people with fathers in manual jobs. Runciman's data indicate that the proportion of the downwardly mobile in the manual strata supporting the Labour Party is hardly higher than the proportion of people in the white-collar strata. McKenzie & Silver (1968) reported that the inter-generation downwardly mobile members of their working-class sample were twice as likely as the non-mobile members to vote Conservative, though the downwardly mobile were under-represented in their limited sample. Also of relevance is the finding of Goldthorpe *et al.* (1968) that the most potent social factor increasing support for the Conservative Party in their sample of 'affluent' male manual workers was 'white-collar affiliations' through either the family origins or job histories of the men or their wives. If we assume that the downwardly mobile are twice as likely as their non-mobile peers to support the Conservative Party, given the high rate of downward mobility into the manual strata (Stacey, 1968), it would appear that the downwardly mobile are responsible for about 25–30 per cent of the working-class Conservative vote and make a substantial contribution to the maintenance of working-class Conservatism.

Though the attitudes of working-class Conservative and Labour supporters converge on some issues they differ substantially in relation to the political parties and the power system in British society (Butler & Stokes, 1969; Goldthorpe *et al.*, 1968; McKenzie & Silver, 1968; Nordlinger, 1967; Runciman, 1966). The Conservative supporters view the parties in terms of policies, leadership qualities and what Conservatives conceive as the 'national interest', while Labour supporters see the parties primarily in terms of class interests. For the latter, the trade unions and the Labour Party are the means by which they can defend and promote their own material interests. For the militant minority of Labour supporters they are also seen as possible instruments for radically changing society. Though antipathy to both big business and the trade unions is widespread in the manual strata, there is a marked relationship between positive or critical attitudes concerning the power of big business and the trade unions, and party loyalties in the direction one would expect. Conservative supporters are far less likely than Labour supporters to see the upper class as hostile to working-class interests. McKenzie & Silver (1968) found that Conservative and Labour supporters disagreed more sharply in their perception of the upper-class attitude to the working class than on any other issue examined in their study. Abrams (1968) concluded that his recent survey findings suggest: 'although the working class Conservative may be "spurious" in the sense that he has not really been assimilated within the middle class (or the Conservative Party), yet he has learned and accepted the "correct" conventional middle class responses when called upon to assess the character and purposes of the British nation and its people'.

According to the *embourgeoisement* explanation of working-class Conservatism, as manual workers and their wives become more prosperous they (1) increasingly assume 'middle-class' values, attitudes, consumption habits, social perceptions and political allegiances, and (2) become progressively assimilated into 'middle-class' society (Abrams, 1960, 1964; Butler & Rose, 1960; Lipset, 1960*a*; Zweig, 1961). This explanation was widely accepted before Labour's 1966 election victory, even though it was subjected to effective criticism (Goldthorpe & Lockwood, 1963; Lockwood, 1960). The idea that 'affluent' manual workers have become or are becoming middle class is still frequently stated. Leaving aside the issue as to whether middle-class society is prepared to assimilate even the prosperous manual workers and their dependants, the *embourgeoisement* thesis is not well supported by the evidence collected from Runciman's national sample (Runciman, 1964) and the smaller samples of Goldthorpe *et al.* (1968), McKenzie & Silver (1968) and Nordlinger (1967). The Goldthorpe *et al.* study was designed as a critical case analysis of the thesis and its authors came to the conclusion that 'no process of *embourgeoisement* is evident among our sample of affluent manual workers', though it must be added that the authors never define the key term and the design of this study leaves much to be desired. One further critical point, the term *embourgeoisement* is an unfortunate choice because this explanation focuses upon the lower as much as the upper sections of the white collar strata. It seems odd to treat clerks, draughtsmen, salesmen, technicians, nurses, school teachers and occupationally similar groups as bourgeois, given the use of the term by such influential writers as Marx, Engels, Weber, Durkheim, Tawney and Sorokin.

It could be argued that the association between self-rated middle-class status and Conservatism in the manual working class provides some support for the *embourgeoisement* explanation if this self-rating involves identification with the middle class in the white-collar strata seen as a normative reference group. Nordlinger's

(1967) empirical analysis of the relationship between middle-class self-rating and Conservative voting strongly suggests that manual 'middle-class subjectives' are economically content people dissociating themselves from low status and the unrespectable working class. They display little concern with becoming *bona fide* members of the middle class and use it as a reference group only in a very limited way. If McKenzie & Silver's (1968) finding that home-ownership, personal savings and middle-class self-rating are particularly associated with Conservative voting among people with incomes *below* the working-class average is verified (pp. 96–97), then we have a direct refutation of one form of the *embourgeoisement* thesis. With regard to middle-class self-rating, the data presented by Runciman (1966) show that the manual middle-class subjectives in the lowest of income groups most frequently support the Conservative Party. In the panel study of Butler & Stokes (1969) it was reported that the small group of people in their national sample who moved from council housing to owner-occupation (mortgage ?) between the summer of 1963 and the autumn of 1964 displayed a tendency to change their self-rating from working to middle class *and* to become more Labour in their partisanship, despite the trend to the Conservatives in the country as a whole during this period (pp. 101–104). At the very least these data do not support the *embourgeoisement* thesis.

The social pressure theory seeks to explain working class politics through the pressures exerted on individuals and groups by the mass media, employment and the social environment (Goldthorpe *et al.*, 1968; Ingham, 1969; Miliband, 1969; Parkin, 1967). The dominant institutions of British society, with the exception of the trade union and co-operative movements, are hostile to or out of accord with the ideology of Socialism and in harmony with the ideology of Conservatism. Parkin (1967) put forward the case that 'the Labour Party, or any Socialist Party in a capitalist society, is not in contest simply with the Conservative Party, or its equivalent elsewhere, but is in a sense at odds with the dominant institutional orders and central values of the society—of which the Conservatives may be said to be the political guardians'. According to this line of argument, given the conservative influence of the Established Church, the monarchy and aristocracy, and the armed forces, the police, the educational system, the mass media and the institutional complex of capitalism, and the material resources of the Conservative Party, it is to be expected that Conservative values, attitudes and voting will be the common outcome in all classes. Left-wing attitudes and support for Left parties are produced and sustained in subcultures whose norms and values are in opposition to, or at variance with, those of the dominant institutions and higher strata of society. At the same time these subcultures provide necessary buttresses against the established order. Large factories and workshops, trade unions, some nonconformist churches, isolated communities and homogeneous working-class communities have provided the conditions for the dissemination of collectivist values and the growth of a radical working-class outlook largely committed to the Labour Party. It follows that working-class Conservatism can be expected to be strongest where involvement in such subcultures is weakest and where, as a consequence, there is more exposure to the dominant value system of capitalism.

Miliband (1969) argues that Labour leaders have through the century consistently succumbed to conservative pressures, particularly with regard to the definition of the 'national interest', and turned themselves into propagandists against many Socialist ideas and purposes. Hence, he concludes, they have contributed in a major way to the 'attunement' of large parts of the manual working class to Conservative ideology. In an earlier analysis he also laid stress on the 'dismal record' of the Labour Party in government, diagnosed primarily as the outcome of Conservative influences (Miliband, 1967).

To date the social pressure theory has only been presented in a general way, and at present it does no more than provide a framework for the analysis of political attitudes and behaviour. The empirical findings relating to (1) the political consequences of downward mobility (Stacey, 1966) and white-collar affiliations (Goldthorpe *et al.*, 1968; Ingham, 1969); (2) the negative association between increasing plant size and the extent of right-wing attitudes and voting (Ingham, 1969); (3) the slightly higher rates of Conservative support found amongst working-class women compared with working-class men (Parkin, 1967); (4) the concept of political generations (Hampton, 1969); and (5) the positive association between the increasing social heterogeneity of a community and the extent of right-wing attitudes and voting (Parkin, 1967) have been shown to be consistent with the social pressure theory. Though it is supported by the relevant evidence collected in recent surveys, these surveys have all been based on relatively simple research designs usually involving no more than descriptive interpretation of the data obtained. Such data analyses are inadequate to test the theory rigorously because they cannot methodically deal with the interactions of the variables studied. For example, there exists a difference of opinion as to whether trade-union membership has an *independent* effect on attitudes and voting (Butler & Stokes, 1969; Goldthorpe *et al.*, 1968; Hampton, 1969; Ingham, 1969; McKenzie & Silver, 1968). There is a need for the social pressure theory to be specified in a precise way and empirically tested by studies using the methods of multivariate analysis (e.g. see Gullahorn, 1967; Hope, 1968).

Nordlinger (1967) and McKenzie & Silver (1968) agree that working-class Conservative supporters can be divided into two types: deferentials and non-deferentials. The deferentials attribute superior leadership abilities to people of high social status and, therefore, prefer aristocratic and upper-class leaders. They endow the Conservative Party leaders with intelligence, skills and virtues beyond the capacities of any critics or opponents. They see these leaders as uniquely qualified by birth, training and experience to govern the nation, and also as their political benefactors. The non-deferentials, though usually holding some deferential attitudes, support the Conservative Party on more pragmatic grounds. They believe that Conservative policies are to the advantage of the working class (and other classes). They see leadership as open to attainment by talented individuals of ordinary background. But they are weaker supporters of the Conservative Party than deferentials, less frequently rate themselves middle class, and are also more likely to make self-interested political demands on the Conservative Party. The evolution of these two types of Conservative supporter is explained largely by (1) the conservative influence of the educational system and the mass media; (2) the success of the Conservative Party's political appeal to the working class since the 1867 Reform Act; particularly its claims to be the party best fitted for government and the defence of 'national interests', and to guarantee the standard of living of the working class; and (3) the relative insensitivity of first the Liberal leadership, then the Labour leadership to the working-class section of the electorate, the support of a majority of working-class people being more or less taken for granted.

On the basis of similar items Nordlinger (1967) placed 28 per cent, and McKenzie & Silver (1968) 26 per cent of the Conservative supporters in their samples in the deference category. They also placed 13 and 20 per cent respectively of the Labour supporters in this category. Though it would appear from these percentages that social deference only explains at best a minority of cases of working-class Conservatism, these investigators believe otherwise. Nordlinger argues that deference is of crucial importance because the 'political culture' of the working class is distinctively characterized by deferential attitudes to authority. McKenzie & Silver argue that deference leads a clear majority of working-class people into an 'acquiescent acceptance' of the class structure and their own subordinate place in it. Runciman's (1966) data on the reasons for Conservative Party support presented by working-class Conservatives (pp. 180, 186) suggest that deference makes a significant contribution to working-class Conservatism, but that it is hardly more important than what Runciman calls 'an acquiescence of tradition and habit' which may depend on family or friends or local custom. No data relevant to the deference typology were published for working-class Labour supporters.

The deference typology presented by Nordlinger (1967) and McKenzie & Silver (1968) is open to criticism on at least two grounds. Firstly, there is no evidence that their deference and non-deference items hang together independently. Neither cluster nor factor analysis was utilized in these studies. Secondly, there is no evidence for the reliability or objective validity of these items, only total reliance on face validity. It could be argued that their data indicate the existence of two attitudinal dimensions rather than two personality types. Further, their results could be interpreted as suggesting that working-class Conservatives (1) tend to varying degrees to be power-oriented and authority-conscious as well as, or rather than, deferential; and (2) are pragmatic to the point of believing that principles are of little account if they count at all. Such an interpretation would not be out of accord with what is known about the psychological characteristics of white-collar Conservative supporters (Stacey & Green, 1968). To explore further the organization of attitudes in the working class the following study was carried out.

METHOD

In a previous study the present writers developed a 58 item questionnaire which established the presence of firm attitudinal distinctions, including social deference, between male white-collar supporters of the Conservative and Labour Parties (Stacey & Green, 1968). As inter-generation downward mobility, white-collar affiliations and middle-class self-rating all play a part in working-class Conservatism it was decided to use these 58 items to investigate the attitudes of manual working-class males. These items, together with 11 items derived from questions used to elicit information about each respondent's occupation, age, income, etc., are described in the appendix.

The questionnaire was distributed among the public at large. The sample as approached consisted of 554 men in manual occupations who lived in Birmingham, Coventry, Glasgow and London. They were the male working-class members of a large probability sample selected randomly from electoral registers. Each man was asked at his home to co-operate in a social survey. If he agreed a questionnaire and a stamped addressed envelope were left with him. He completed the questionnaire in his own time and posted it back to one of the writers. The procedure was anonymous. Some 333 questionnaires were returned and 302 were used in the analysis. The remaining 31 were from SNP supporters, Liberals, Communists and non-voters, each group being too small to warrant analysis. The lowest proportional return was from men in the over-60 age-group.

For each of the 121 Conservative and 181 Labour respondents of the study there were 69 variables, comprising 58 attitude items and 11 descriptive characteristics (marker variables). Using these variables a 69 × 69 correlation matrix was obtained. Three factor analyses, with Varimax rotation to simple structure, were carried out. At the first level 23 primary factors with a latent root greater than unity were extracted. The 11 marker variables were correlated with the 23 factors. At the second level seven factors were extracted and at the third level three factors. At both levels the marker variables were correlated with the factors extracted.

RESULTS

Loadings on 14 of the 23 first-level factors and the correlations of the 11 marker variables with these factors are given in Table 1. None of the marker variables correlated significantly with any of the nine remaining first level factors and these are omitted from Table 1. Support for the Conservative Party is positively associated with three factors at a level higher than 0·20: I (0·44), VIII (0·35) and XXIII (0·37).* Factor I has high loadings on variables 42, 43, 45, 47, 51, 52 and 55, and is more characteristic of older than younger men. It is a royalist factor involving allegiance to the throne and respect for British traditions. Factor VIII has its highest loadings on variables 15, 21, 29 and 59, and is especially associated with older men. It is a factor of firmness in discipline involving support for the use of corporal punishment. Factor XXIII has its highest loadings on variables 24, 46, 48 and 50. This is a factor of unswerving loyalty to authority and leadership which is applied particularly to 'born' leaders.

Support for Labour is positively associated with five of these factors: II (0·20), IV, (0·22), XVI (0·34), XVIII (0·53) and XXII (0·25). Factor II has high loadings on variables 3, 7, 11, 16, 37, 39, 47 and 69. This is a factor of sympathy and sensitivity towards others and is associated with an expressed fondness for the father. Factor IV has its highest loadings on variables 8, 13 and 20. It combines a belief that criminals should be reformed rather than punished, with opposition to flogging and capital punishment. Factor XVI has high loadings on variables 18, 40, 58, 61 and 67. It relates respondent support for Labour to a paternal support for that Party which is associated with belief in God and a self-help philosophy. Factor XVIII has high loadings on variables 11, 18, 33, 42, 44, 64, 65 and 66. This is an egalitarian factor more pronounced amongst those with a middle-class than working-class self-image (but not related to inter-generation downward mobility). Factor XXII has its highest loadings on variables 23, 38 and 48. It is a critical, questioning factor associated with an expressed dislike of the father.

---

* A negative sign means the association is with the right of the item.

Table 1. *Primary factor loadings: first level*

	Factors													
Variable	I	II	III	IV	V	VII	VIII	XV	XVI	XVII	XVIII	XIX	XXII	XXIII
1	+03	+26	**−31**	−10	−05	−09	−21	−11	+10	−14	+14	−05	−22	+05
2	−15	+05	+04	−04	−03	−05	−07	−03	+08	00	+02	−08	+30	+23
3	−09	**−66**	+15	+05	+03	00	−11	+16	+15	−14	−07	−03	−17	+10
4	−08	+11	+03	+19	−02	+15	+08	**−89**	−01	+15	−05	+03	+16	00
5	−04	−02	−01	−08	−06	+05	−21	+01	+09	+10	−04	−13	−16	+01
6	−14	+08	+22	+06	−15	+11	+13	−04	00	−02	+04	−13	−15	+09
7	+09	**+42**	+12	+20	+04	−01	+25	+17	+14	+10	+03	−13	00	+02
8	−05	−06	−03	**−80**	+03	−04	−04	−01	−09	−02	−14	+01	−01	−20
9	−04	−15	+13	+05	+06	−11	+04	**−1.0**	+02	+26	−08	+08	−14	−17
10	+03	+02	−09	+18	+09	−18	+02	−02	+07	+18	+01	00	+13	−08
11	+04	**−40**	+04	−05	−08	+05	+03	−02	−07	+12	**−34**	+13	−12	−20
12	−18	+04	+18	+01	+01	−04	+11	−17	−03	−11	00	+14	−21	−07
13	−06	−02	−12	**−84**	−03	−06	+25	+15	+11	+02	−12	−05	−05	−07
14	−07	+05	+01	−02	+10	−10	+04	−03	+03	+16	+07	+11	−24	+21
15	−03	+25	+05	−28	−08	+08	**−42**	−19	−09	+07	+07	−13	+09	+23
16	−04	**−35**	−05	−18	−08	**+48**	−18	+24	−09	+16	+01	+10	−14	−24
17	+03	+01	+20	+04	−12	−23	−15	−01	−23	−05	+09	+18	00	+06
18	−08	+12	−25	+08	−06	+06	−10	+11	**−31**	−05	**+30**	−23	+02	−11
19	−07	+08	−05	+12	−03	+15	−03	+04	−01	+01	−07	−09	−04	−13
20	−02	+15	+01	**−81**	−05	−03	+07	+21	−09	+12	−10	+07	+09	+02
21	−07	−04	+07	+21	−17	+28	**−78**	+06	−04	−01	−17	−23	−26	−01
22	+02	−08	−23	+01	00	−13	+07	+12	−11	+06	−19	−08	+13	+04
23	+08	−10	−06	−18	+19	+11	+04	+07	−16	−18	+03	−06	−16	+18
24	−02	+02	−04	−07	−10	−08	+11	−15	−18	−12	−19	+05	**−39**	**−95**
25	+05	−16	−03	−31	−08	+05	+19	−18	+12	+02	+16	+05	−06	+11
26	+10	+02	**+34**	−12	+09	+20	+27	−05	−19	−03	+10	−28	+19	−07
27	−07	+03	−10	+02	+03	−03	−24	−20	+10	−17	00	+11	−07	−04
28	−14	+15	**+41**	−20	+03	+17	+05	+11	+18	+14	−08	−19	−13	+14
29	**+31**	+06	+03	+07	+17	+23	**−89**	−01	−05	+13	+05	−10	−11	00

Table 1—*continued*

Variable	Factors													
	I	II	III	IV	V	VII	VIII	XV	XVI	XVII	XVIII	XIX	XXII	XXIII
30	+07	00	**+85**	+12	+04	−04	−13	−15	00	+18	+02	+03	−12	+04
31	−03	+05	00	+12	+13	+23	−04	00	+09	−11	+22	+24	−05	+04
32	+03	+02	−02	−02	+05	+22	−07	−27	+12	00	+01	+10	+16	−21
33	+12	−13	−11	−10	+08	**−30**	+02	−02	−21	−03	**−57**	−02	+03	−02
34	+15	+05	**−42**	−04	−14	+08	+02	−11	+05	+04	−24	−03	+06	+09
35	−03	−23	+20	−09	−07	−20	+01	−20	+18	−20	−16	+06	+17	+02
36	+04	−16	+12	+05	−19	+26	−24	+09	−25	−14	+06	+20	+27	+06
37	+06	**+73**	+07	−01	−07	−12	−22	+11	−08	+05	+16	−02	−13	+06
38	+08	−18	+07	+08	+09	−07	+05	−08	−15	−05	−06	−28	**−39**	−09
39	+07	**+61**	+11	−03	−13	−11	+08	+04	−16	−09	+03	+17	−12	−28
40	−11	+13	−04	+10	−14	−05	+03	−11	**−31**	−01	−20	−04	−07	+08
41	−08	+06	+02	+06	−02	+01	+02	+15	−05	+17	+08	+04	+04	+21
42	**+79**	+06	+10	+12	−15	−02	−10	+04	−14	00	**−30**	+09	+15	−17
43	**−57**	−13	−21	+12	00	00	+01	−12	−01	+09	−16	+10	+06	+01
44	+12	−14	−02	−14	+06	+11	−08	−11	+11	+06	**−81**	**−33**	+16	−15
45	**−88**	−02	+10	−03	+01	+11	+13	−02	−16	−06	+10	+07	+05	−12
46	−09	+16	−05	−19	+02	−09	−15	+02	−02	−04	+17	+04	−20	**−72**
47	**+38**	**+44**	+11	−06	−08	−09	−03	−10	−03	00	−16	−10	−03	−11
48	+08	+06	+11	00	+05	−05	−23	+02	+16	+09	+13	−04	**−1.0**	**−36**
49	+04	−04	+02	−09	−01	+03	−05	−04	−03	−01	−04	−05	+17	+04
50	+04	+05	−07	+05	−08	+06	+05	−01	+01	+08	+08	+27	−13	**+35**
51	**−42**	−02	+02	−01	−01	−19	+07	−01	−08	−08	+18	−05	+14	−28
52	**+93**	+06	−03	+05	+01	−19	−16	+05	+14	−03	00	−01	−14	+06
53	00	−02	+04	+09	−13	−12	+16	+23	−07	−01	+09	+18	−03	+12
54	−13	−04	+04	+28	+04	−14	00	+16	−01	+01	−10	−29	+01	−12
55	**+34**	+01	−04	+10	+11	+09	+11	+04	−14	+02	00	+16	−02	+10
56	−19	+03	+02	−08	+04	−22	−12	+13	−14	+15	−08	−19	+12	−05
57	+08	−02	+08	+01	00	−02	+13	−14	+08	−18	−15	+05	+23	+20
58	−08	+07	+12	−01	00	−16	−16	+01	**+31**	+05	−11	00	−20	−02
59	−04	+02	+17	−05	+14	−08	**−53**	**+32**	−01	−04	−07	+22	−05	+10
60	+08	−02	+11	−07	+07	−05	−17	−25	−29	**+90**	+04	−04	−08	+19

Table 1—*continued*

Variable	I	II	III	IV	V	VII	VIII	XV	XVI	XVII	XVIII	XIX	XXII	XXIII
	Factors													
61	+01	+01	00	+02	−05	+24	−16	+08	**+41**	+01	−24	**−1.0**	−06	+04
62	+05	−11	−17	+02	+08	−14	+02	+20	+11	**−89**	−03	−02	+07	−06
63	+02	−26	+04	+11	**+86**	−02	−02	−08	−16	−08	+03	−01	−13	+03
64	−01	−15	−04	00	−26	−07	+10	+03	−02	**+46**	**+52**	+08	−03	+03
65	+14	−11	−08	+07	**−77**	+03	+13	−05	+22	−04	**+39**	−09	−04	−11
66	**+30**	−15	+02	−11	+04	+08	+04	+12	−13	+02	**−44**	+01	−01	+11
67	−10	+17	−01	−04	+20	+02	−06	+01	**−1·0**	+23	−04	**+40**	+15	−17
68	−11	−14	+03	+09	−04	**+98**	−29	−05	+03	+07	+02	−27	+08	+17
69	−16	**+35**	−17	+03	+13	**+48**	+01	−05	−25	−07	−13	00	−06	−02

Correlations of marker variables with primary factors

Variable	I	II	III	IV	V	VII	VIII	XV	XVI	XVII	XVIII	XIX	XXII	XXIII
59 Age	**−30**	−02	**+20**	+10	+08	**−24**	**−43**	**+35**	+18	+06	−11	**+40**	+02	+01
60 Income	+11	−01	+05	−07	+03	+08	−04	−13	−11	**+48**	+11	−06	−05	+13
61 God	+15	+13	−03	−06	−01	+05	+04	−12	−02	−08	−03	**−61**	−05	+08
62 Occ. status	−01	−03	−13	−02	+05	−18	−11	−07	−06	**−64**	+02	−10	−03	+01
63 F's status	+04	−10	+02	+08	**+73**	+07	+05	00	−07	−07	+07	−03	−07	+03
64 Class	−15	−12	−07	+01	**−20**	−14	−12	+01	+16	**+34**	**+46**	+08	+14	−18
65 F's class	+04	−11	−08	−04	**−64**	−10	+03	−13	+01	−02	**+27**	−10	−02	−03
66 Conservative	**+44**	**−20**	+12	**−22**	−08	+17	**+35**	+08	**−34**	−07	**−53**	−05	**−25**	**+37**
67 F. Conservative	+02	−01	+03	+01	+06	+05	+03	+03	**−61**	+03	−07	+02	−07	+05
68 F. strict	+13	+06	+10	−01	+04	**+65**	+07	−08	−05	+01	00	−04	−06	+13
69 Fond of F.	+05	**+41**	−11	+01	+17	**+47**	+08	−24	−12	**−20**	−12	+08	**−26**	+16

Table 2. *Primary factor loadings: second level*

Variable	Factors I	II	III	IV	V	VI	VII
1	+02	**+68**	−09	−04	−16	−07	+03
2	−32	+16	+19	+14	+16	+05	−09
3	−11	−17	−29	+21	+16	00	−13
4	+11	**+33**	+13	+01	−03	00	−03
5	+07	−12	+02	+09	+09	−07	−29
6	−07	+09	−31	+24	+12	+10	−02
7	+09	−12	**+73**	+10	+16	−03	+07
8	−01	+04	**−54**	−13	+13	+06	−12
9	+17	+21	+13	−12	−01	−08	−15
10	−35	+07	+11	+01	+16	−19	+17
11	−11	**−58**	−13	+20	+13	+22	−06
12	−09	+18	+35	−04	+05	+03	−17
13	+18	−04	**−37**	−10	+06	+07	−05
14	+27	−17	−09	−20	**+36**	−14	+06
15	−17	+29	+08	+05	00	+11	**−39**
16	+13	00	−16	+03	+29	−22	**+32**
17	−12	−11	+23	+09	−01	+10	**−50**
18	+01	+03	+34	−07	−17	+06	**−42**
19	+30	−02	**+57**	00	−02	−17	−08
20	+17	−09	**−38**	+05	+09	+04	−06
21	−06	+02	+13	+05	−03	−02	−18
22	+15	+01	+17	+02	+03	−07	−11
23	−06	+08	−16	+01	**+34**	+06	**−47**
24	−01	−21	+09	−05	−01	−10	−15
25	−09	+18	−28	+09	**+36**	+22	−01
26	+04	+06	−03	−18	**+50**	+26	−13
27	−14	−11	+07	+06	+15	+12	+03
28	+01	−16	−01	+17	**+56**	+10	+06
29	−12	+23	+05	00	−01	−15	+05
30	−02	−14	−07	+16	+33	−04	−12
31	−18	+11	**+45**	+06	+10	**−30**	−09
32	−01	**+38**	+10	−01	+15	**−63**	00
33	+34	−16	−11	+14	−09	+24	−11
34	−06	+10	+33	−23	−05	+15	−06
35	+21	+10	**−43**	+24	−19	−27	−27
36	−04	−06	**−50**	−03	**+37**	−14	−05
37	+04	+22	**+40**	+10	−06	−08	−14
38	+25	+29	−13	+17	−09	**+46**	−04
39	+05	00	**+41**	+10	+30	00	−09
40	**+58**	+05	−02	+16	−17	−03	**+31**
41	+15	+13	−15	+03	−05	+03	+27
42	**+45**	−17	+12	−08	+23	+09	00
43	**−48**	−16	+12	+21	−13	+02	+10
44	**+40**	−11	−09	−04	−01	−09	+05
45	**−53**	−07	−11	+11	+01	−18	−07
46	−16	+12	+20	−02	**−31**	**−30**	+05
47	**+56**	+20	+06	+03	+16	−11	−08
48	+02	+02	+19	+21	+03	+05	−11
49	**+42**	+01	−03	+04	+06	−01	+11
50	+03	00	+16	**+69**	+16	+21	−04
51	−33	−09	−13	−28	−04	**−32**	−18
52	**+57**	+13	+08	−06	+04	+07	00

Table 3. *Primary factor loadings: third level*

Factors I	II	III
−42	−19	−02
−38	+12	+08
+18	+17	+28
−33	−18	−05
−07	+11	+04
+08	−05	+31
−39	−05	−04
+19	−10	+11
−31	−13	−14
−22	+28	+19
+51	+13	+12
−47	−02	−17
+24	−19	+05
+03	−17	+18
−43	+01	−15
+05	−06	**+45**
−23	+17	−24
−30	+06	**−43**
−36	−04	−16
+27	−11	+19
−16	+11	−08
−16	−03	−04
−32	−04	+10
+03	+20	−11
−11	−27	**+38**
−28	**−42**	+08
00	00	+07
−08	−17	**+46**
−24	+11	−05
−06	+06	+31
−53	+32	+03
−58	+32	+29
+30	−29	−05
−28	−14	**−37**
+15	+21	+19
+11	+06	**+40**
−46	+01	−11
+02	**−58**	−04
−42	−09	+10
+02	−24	−07
+14	−22	+13
+02	**−40**	+03
+12	**+44**	+02
+19	−17	+05
−01	**+57**	+13
−13	**+39**	−16
−26	**−40**	+14
−17	00	+05
+07	−32	+13
−10	−09	**+45**
−03	**+52**	−15
−09	**−52**	−04

Table 2—*continued*

Variable	Factors I	II	III	IV	V	VI	VII
53	−31	+19	−23	−05	+13	+01	+15
54	**−40**	−05	+05	+10	+19	−04	−13
55	**+42**	**+37**	+07	+15	00	+04	+29
56	**−51**	+17	−33	−13	−18	+04	+09
57	**+51**	+17	+06	+03	+07	−04	+06
58	−29	+19	−05	+16	00	**−39**	−01
59	−19	−24	+02	**+42**	−17	−13	−04
60	−04	+11	+01	−04	+08	+02	+12
61	+12	+13	+04	−17	+02	+14	−15
62	−04	+19	−15	00	−16	+07	**−30**
63	−25	+19	+03	+03	−01	+12	+21
64	+07	−02	−01	−11	−14	**−30**	−05
65	+13	−03	−09	−04	−06	+01	**−30**
66	+22	−10	−16	+13	+16	**+34**	+08
67	−01	−02	−04	+01	+04	+21	−09
68	+05	+04	+12	+01	**+36**	+01	+12
69	+02	**+38**	+10	+22	+08	+01	+15

Correlations of marker variables with primary factors

Variable	I	II	III	IV	V	VI	VII
59	**−22**	**−31**	−07	**+42**	**−25**	**−28**	−01
60	−01	+11	+03	−04	+12	+04	+10
61	+14	+20	+12	**−21**	+04	+18	−16
62	−03	+16	−02	−04	−14	+03	**−31**
63	−19	+16	+07	00	+03	+09	+18
64	−06	−06	+04	−08	−14	**−27**	−12
65	+11	−02	−05	−06	−09	+01	**−27**
66	**+38**	−09	**−28**	+15	+19	**+42**	**+24**
67	+06	+01	−05	−02	+03	+19	−04
68	+08	+14	+08	+01	**+38**	+04	+14
69	+07	**+40**	+15	**+21**	+18	+05	+11

Table 3—*continued*

Factors I	II	III
−06	+05	+18
−19	+32	+13
−15	**−50**	+20
+12	+27	−08
−13	**−42**	+10
−24	**+48**	+26
+23	**+48**	+18
−09	−09	+06
−16	−25	−21
−09	00	−17
−09	−03	+04
00	+23	−08
+02	−01	−15
+24	**−43**	+16
+03	−14	−05
−23	−19	+24
−33	−19	+25

I	II	III
+20	**+47**	+15
−07	−09	+06
−17	**−20**	**−20**
−12	+05	−19
−08	−03	+03
−04	+24	−11
−01	+01	−14
+32	**−48**	**+26**
+04	−14	−03
−16	−20	**+23**
−26	−18	**+21**

The remaining six factors in Table 1 are clearly interpretable in terms of item content and are related to the marker variables age, income, occupational status, occupational status of father, and strictness of father.

The second-level factor analysis adds particularly to our analysis of working-class Conservatism (Table 2). Support for the Conservative Party is positively associated with three of the seven factors at a level higher than 0·20: I (0·38), VI (0·42) and VII (0·24). Factor I has high loadings on variables 40, 42, 43, 44, 45, 47, 49, 52, 54, 55, 56 and 57. It is a factor of 'patriotic' support for Britain, British traditions, royalty and 'natural' rights. Factor VI has high loadings on variables 31, 32, 38, 46, 51, 58, 64 and 66 and is associated with older men and a middle-class self-rating. It is a mixture of elitism and 'natural' inegalitarianism, self-reliance and national solidarity. Factor VII has high loadings on variable 15, 16, 17, 18, 23, 40, 62 and 65. This is a factor of self-interest combined with a certain firmness of conviction and toughness. It is associated with non-skilled occupational status and assignment of the father to the middle class (an assignment not related to the father's income or occupational status).

Psyl. RESEARCH
R
3.3.81

Supporting Labour is related to factor III (0·28). It has high loadings on variables 7, 8, 13, 19, 20, 31, 35, 36, 37 and 39. This factor has three components: egalitarianism, sympathy for those needing help, including the poorer countries, and a reformist attitude towards criminals. The remaining three second-level factors are related to age, father's strictness and fondness for the father.

The third-level factor analysis produced three factors (Table 3). Conservative Party support is related to factor I (0·32) and factor III (0·26), Labour Party support to factor II (0·48). Factor I has high loadings on variables 1, 2, 7, 11, 12, 15, 19, 31, 32, 37 and 39. It is a complex factor involving self-interest, inegalitarianism, lack of sympathy for the underdog, emphasis on personal initiative, an elite leadership orientation and a nationalistic outlook. Factor III has its highest loadings on variables 16, 18, 25, 28, 34, 36 and 50. This is a factor of discipline, toughness and submission to authority. It is associated with a strict father and fondness for the father. Factor II has high loadings on variables 26, 38, 42, 43, 45, 46, 47, 51, 52, 55, 57, 58 and 59. It is an anti-royalty, anti-aristocracy factor involving (1) opposition to hereditary privilege, patriotism and many British traditions, and (2) an emphasis on the current state of affairs.

Table 4. *Primary factor loadings*

	Factors			
	1st level		2nd level	3rd level
Variable	I	XVIII	VI	II
11		−34		
18		+30		
26				−42
31			−30	
32			−63	
33		−57		
38			+46	−58
42	+79	−30		−40
43	−57			+44
44		−81		
45	−88			+57
46			−30	+39
47	+38			−40
51	−42		−32	+52
52	+93			−52
55	+34			−50
57				−42
58			−39	+48
59				+48
64		+52	−30	
65		+39		
Correlations of marker variable 66 *with these factors*				
	+44	−53	+42	−48

Four factors seem to be of major importance in view of their relationship with item 66. These are: first level/factor I—pro royalty, aristocracy and British traditions; first level/factor XVIII—egalitarianism; second level/factor VI—elitism, opposition to egalitarianism, self-reliance; third level/factor II—opposition to royalty, aristocracy, patriotism and outmoded traditions. Table 4 gives the highest loading variables on each of these factors.

DISCUSSION

Though the manual working class comprises almost two-thirds of the population, it is a relatively disadvantaged social formation which is, and always has been, highly diversified. The influences that account for Conservatism (and conservatism) in this class are numerous and varied. Inter-generation downward mobility makes a substantial contribution to working-class conservatism and the Conservative vote. Regional, religious, age and sex influences make a significant but small contribution. Runciman (1966) offered the following sociological caricature of the *ideal type* of working-class Conservative supporter: 'a woman in her seventies living in a country district in the Midlands whose father was in a non-manual occupation, who stayed on at school beyond the minimum leaving age, who thinks of herself as "middle-class", and who would like to attend regularly at an Anglican church but is prevented by age or illness from doing so'.

Ideas of class interest and class conflict are far less salient to working-class Conservative than Labour supporters. The former tend to feel that their needs are met by the political system, are economically more satisfied, either lack discontent or if they are discontented, are less likely to channel discontent into collective action and more likely to believe that individualistic solutions are available, are more traditional in their attitudes, aspirations and perception of social relationships. They tend to come from small families and rather more frequently have personal savings. There also appears to be a greater tendency for the children of working-class Conservative supporters to be upwardly mobile into the white-collar strata (Butler & Stokes, 1969; Toomey, 1969). These characteristics suggest a working-class Conservative syndrome of thrift and self-help involving a desire for long-term stability and security. Conservatives may also place a higher value on education as a means of personal advancement.

The differences reported in the previous section between male working-class Conservative and Labour supporters agree reasonably well with the patterning of attitudes revealed in the complementary survey of white-collar males, even to the extent that Conservative supporters regardless of status tend to express fondness for a strict father. Fourteen of the items listed in Table 4 appeared in the corresponding table of the earlier study. But the patterning of attitudes is looser or less definite in the manual than in the white-collar strata, and the attitudinal differences are less clear cut in the former strata. It seems that both the Conservative and Labour Parties attract working-class support for a good many of the reasons they attract white-collar support.

Psychoanalytic speculations apart, the working-class Conservative supporter can be most readily distinguished from his Labour non-comrade by his class imagery, class identification, his attitudes towards traditions, authority, and dominant institutions and symbols, combined with support for, or at least acceptance of elitism, hierarchy and privilege. Rather than seeing the social structure in terms of class interests, he construes it as a reasonable hierarchy with himself looking up to his 'betters' and down on his 'inferiors'. This construing accords well with the Conservative Party's projection of itself as a national party representing all interests within an integrated if stratified society. The 'us' and 'them', 'haves' and 'have-nots', 'bosses' and 'workers' type of dichotomization common among working-class Labour supporters reveals a radically different outlook even when it is accompanied by some conservative values and beliefs.

It is possible to interpret the psychological characteristics of working-class Conservative supporters as a deferential reaction to the elite in the social order and their voting behaviour as the means by which they register support for the party of the upper class. It also appears that they see the answers to any and every social problem as being provided by leaders from the upper class and the professions. But it is not legitimate to interpret these characteristics as indicating that social deference is the distinctive feature of this working-class group. Social deference is at least as widespread among white-collar Conservative supporters who average higher levels of education, income and status (Butler & Stokes, 1969; Stacey & Green, 1968). Deference to those in high places, of elite social origin and in positions of authority is common among these white-collar supporters. They also widely believe that the education and business experience of Conservative leaders give these leaders greater general skill in governing, especially in the management of the country's financial and foreign affairs.

It is also possible to interpret the psychological data on working-class Conservative supporters as reflecting a consciousness of authority and an orientation towards the power and influence of the upper strata among reasonably satisfied people deficient or relatively weak in working-class consciousness. The distinct class influence is that, being manual working class, they see their best interests served by following and supporting those in authority/power rather than aspiring for success or to attain authority/power. Such an interpretation encompasses the economic interest variable and the acquiescence in tradition and habit stressed by Runciman (1966). It allows for the continuing evolution of working-class life with changing circumstances and standards, without the necessity of resorting to notions of mass *embourgeoisement* because of rising consumption levels. It is consistent with Parkin's (1967) argument that the Conservative Party's alignment with the dominant religious, educational, social and economic institutions benefits it to some considerable extent in all social strata.

One final issue: the results of this study provide some evidence supporting Miller & Riessman's (1961) critique of Lipset's (1960*b*) analysis of working-class authoritarianism with its emphasis on the supposed authoritarianism of the Left in this class. Brown's (1965) conclusion that no one so far has shown there is an authoritarian of the Left still stands (see also Newton, 1969). Lipset (1960*b*) has also argued that the manual working class is more authoritarian than other classes. But Lipsitz (1965) has shown that when Lipset's data are analysed with controls for education, the greater authoritarianism of the working class disappears. More recently, Form & Rytina

(1969) have severely criticized current views concerning the supposedly high authoritarianism of the poor on the basis of their own and other investigators' evidence.

REFERENCES

ABRAMS, M. (1960). New roots of working class Conservatism. *Encounter* May, 57–59.
ABRAMS, M. (1964). Party politics after the end of ideology. In E. Allardt & Y. Littunen (eds.), *Cleavages, Ideologies and Party Systems*. Helsinki: Westermarck Society.
ABRAMS, M. (1968). Some measurements of social stratification in Britain. In J. A. Jackson (ed.), *Social Stratification*. Cambridge: Cambridge University Press.
ALFORD, R. A. (1964). *Party and Society*. London: Murray.
BROWN, R. (1965). *Social Psychology*. London: Collier-Macmillan.
BUTLER, D. & ROSE, R. (1960). *The British General Election of 1959*. London: Macmillan.
BUTLER, D. & STOKES, D. (1969). *Political Change in Britian*. London: Macmillan.
FORM, W. H. & RYTINA, J. (1969). Ideological beliefs on the distribution of power in the United States. *Am. sociol. Rev.* **34**, 19–31.
GOLDTHORPE, J. H. & LOCKWOOD, D. (1963). Affluence and the British class structure. *Sociol. Rev.* **11**, 133–163.
GOLDTHORPE, J. H., LOCKWOOD, D., BECHHOFER, F. & PLATT, J. (1968). *The Affluent Worker: Political Attitudes and Behaviour*. Cambridge: Cambridge University Press.
GULLAHORN, J. E. (1967). Multivariate approaches in survey data processing. *Multiv. behav. Res.* Monogr. no. 67–1.
HAMPTON, W. (1969). Working-class angels. Society for the Study of Labour History. *Bull. no.* **18**, 54–64.
HOPE, K. (1968). *Methods of Multivariate Analysis*. London: University of London Press.
INGHAM, G. K. (1969). Plant size: political attitudes and behaviour. *Sociol. Rev.* **17**, 235–249.
KELVIN, R. P. (1965). The non-conforming voter. *New Society* 25 November, 8–12.
LIPSET, S. L. (1960*a*). Must Tories always triumph? *Socialist Commentary* November, 11–12.
LIPSET, S. L. (1960*b*). *Political Man*. London: Heinemann.
LIPSET, S. L. (1961). 'Working-class authoritarianism': a reply to Miller and Riessman. *Br. J. Sociol.* **12**, 277–281.
LIPSITZ, L. (1965). Working-class authoritarianism: a re-evaluation. *Am. sociol. Rev.* **30**, 103–109.
LOCKWOOD, D. (1960). The new working class. *Arch. europ. Sociol.* **1**, 248–259.
MACCOBY, E. E. (1954). Youth and political change. *Publ. Opin. Q.* **13**, 23–39.
MCKENZIE, R. & SILVER, A. (1968). *Angels in Marble*. London: Heinemann.
MILIBAND, R. (1967). Labourism and the Labour Government. In R. Miliband & J. Saville (eds.), *Socialist Register*. London: Merlin.
MILIBAND, R. (1969). *The State in Capitalist Society*. London: Weidenfeld & Nicolson.
MILLER, S. M. & RIESSMAN, F. (1961). 'Working-class authoritarianism': a critique of Lipset. *Br. J. Sociol.* **12**, 263–276.
NEWTON, K. (1969). *The Sociology of British Communism*. London: Allen Lane.
NORDLINGER, E. A. (1967). *The Working Class Tories*. London: MacGibbon & Kee.
PARKIN, F. (1967). Working-class Conservatives: a theory of political deviance. *Br. J. Sociol.* **18**, 278–290.
RUNCIMAN, W. G. (1964). 'Embourgeoisement', self-rated class and party preference. *Sociol. Rev.* **12**, 137–154.
RUNCIMAN, W. G. (1966). *Relative Deprivation and Social Justice*. London: Routledge & Kegan Paul.
STACEY, B. G. (1966). Inter-generation mobility and voting. *Publ. Opin. Q.* **30**, 133–139.
STACEY, B. G. (1968). Inter-generation occupational mobility in Britain. *Occup. Psychol.* **42**, 33–48.
STACEY, B. G. & GREEN, R. T. (1968). The psychological bases of political allegiance among white-collar males. *Br. J. soc. clin. Psychol.* **7**, 45–60.

Toomey, D. M. (1969). Home-centred working class parents' attitudes towards their sons' education and careers. *Sociol.* 3, 299–320.

Zweig, F. (1961). *The Worker in an Affluent Society*. London: Heinemann.

Manuscript received 16 January 1970

Revised manuscript received 16 March 1970

APPENDIX

(Items 1 to 58 inclusive are presented in Stacey & Green (1968, pp. 57–60). A duplicated copy of all 69 items is available from B. G. Stacey.)

59. Age: (1) under 30; (2) 30–60; (3) over 60.
60. Income: (1) under £830; (2) 830–£1200; (3) over £1200.
61. Belief in God. Not sure. No belief.
62. Respondent's occupational status as assessed using the M.R.S. scale: C2 D E
63. Father's occupational status as assessed using the M.R.S. scale: A B C1 C2 D E
64. Respondent's self-assignment to a social class: Working class; Middle class; Upper middle class.
65. Respondent's assessment of his father's social class: Working class; Middle class; Upper middle class.
66. Conservative. Labour.
67. Father Conservative. Father Labour.
68. Respondent's estimate of his father's strictness when he was a child: (i) Extremely strict. (ii) Far more strict than most fathers. (iii) A bit on the strict side. (iv) About average. (v) A bit easygoing. (vi) Much more easygoing than most fathers. (vii) Extremely easygoing.
69. Respondent's degree of fondness for his father: (i) I worshipped him. I thought he was wonderful. (ii) I liked him very much. (iii) I liked him a bit more than most other children seemed to like their fathers. (iv) About average. (v) I did not like him quite as much as other children seemed to like their fathers. (vi) I disliked him very much. (vii) I hated him. I thought he was terrible.

Br. J. soc. clin. Psychol. (1971), **10**, pp. 27–29
Printed in Great Britain

Do Spectacles Really Affect Judgements of Intelligence?

By MICHAEL ARGYLE AND ROBERT McHENRY

Department of Experimental Psychology, Oxford University

Video-tapes were made of performers with and without spectacles, seen either static for 15 sec. or being interviewed for 5 min. Judges rated IQs of performers as 12 points higher when they were seen wearing spectacles and were seen briefly; when the performers were seen for 5 min., however, spectacles made no difference. It is argued that this result throws doubt on experiments in person perception using photographs or brief exposure.

In many studies of person perception photographs have been used. This has the advantage that it is possible to do highly controlled experiments in which the effects of particular physical cues can be studied. A possible disadvantage of the method is that the stimuli in question may not have the same effects in real-life conditions where more information is available. It has been found, using photographs, that people wearing spectacles are judged as more intelligent (Thornton, 1943, 1944), and that people with certain facial features are judged as Jewish (Scodel & Austrin, 1957). The hypothesis of the experiment to be described was that the effect of spectacles on judgements of intelligence would disappear if the target person was seen talking for 5 min.

METHOD

Thirty students at a teacher training college, including males and females, were divided into two groups of 15 to act as subjects. Each group watched four video-taped films, each film of a different person, with the object of 'rating personalities'. Taking part in the films were four female performers (A, B, C and D) aged 18–19; A and B were shown on the screen for 15 sec. each sitting still, but C and D were shown being interviewed for 5 min. about their holidays abroad. The first group of subjects saw A without spectacles, B with, C with and D without spectacles; the second group of subjects saw A with, B without, etc. (see Table 2). Two of the target persons (A and C) normally wore spectacles, two did not—to eliminate an alternative possible source of the spectacles effect, that spectacle wearers *without* their glasses look dazed and unintelligent. The independent variable in the experiment was therefore whether or not subjects were seen wearing spectacles. Every other aspect of each target person's performance was the same in both sets of films—including the content of the interviews, which had been carefully rehearsed and standardized. In the 5 min. conversations the target persons were seen being interviewed about their holiday plans. These interviews were fairly relaxed and undemanding, and did not provide clear indications of intellectual ability.

In order to disguise the real purpose of the study, subjects in the experiment were each given a sheet on which various personality dimensions, like *serious–frivolous*, could be evaluated along seven-point scales, and asked to complete it for the person in each film. Among these scales was one which required the subject to draw a cross on a continuum drawn from IQ 100 to IQ 150, with marks at every five points of IQ, according to how they rated the subject's intelligence. It was the latter scores which made up the raw data in this experiment.

RESULTS

The estimates of target person's IQ were analysed by analysis of variance. Interaction between spectacles and length of exposure accounted for most variance ($P < 0.001$).

There was a main effect due to spectacles ($P < 0{\cdot}01$), but it can be seen from Table 2 that this is entirely due to the effect of spectacles with *brief* exposure; under long exposure both wearers of spectacles were actually seen as slightly *less* intelligent. Under brief exposure they were seen as 11·6 to 12·6 points more intelligent.

Table 1. *Analysis of variance: spectacles and time of exposure*

	S.S.	D.F.	M.S.	F	p
Between	62306·50				
Groups (A)	1657·63	1	1657·63	< 1·0	n.s.
Ss within groups	60648·87	28	2166·03		
Within					
Exposure (B)	136·53	1	136·53	1·36	n.s.
A × B	9·64	1	9·64	< 1·0	n.s.
B × Ss within groups	22820·33	28	100·73		
Spectacles	589·63	1	589·63	12·54	< ·01
A × C	333·34	1	333·34	7·09	< ·05
C × Ss within groups	1316·53	28	47·02		
B × C	1267·51	1	1267·51	26·43	< ·001
A × B × C	19·19	1	19·19	< 1·0	n.s.
B × C × Ss within groups	1342·8	28	47·96		

Table 2. *Mean IQ ratings for stimulus persons with and without spectacles t and P values*

Stimulus person	Spectacles	Exposure	Mean rated IQ	t	P
A	no	brief	113·4	3·197	< 0·01
	yes	brief	126·0		
B	yes	brief	128·2	4·187	< 0·001
	no	brief	116·6		
C	yes	prolonged	125·0		n.s.
	no	prolonged	126·3		
D	no	prolonged	122·2		n.s.
	yes	prolonged	119·4		

DISCUSSION

The main hypothesis was upheld—wearing spectacles increases a target person's rated IQ by about 12 points when seen briefly, but has no such effect when he is seen talking for 5 min.

This is a simple and straightforward result, but it has implications beyond the present experiment. A great deal of research on person perception has made use of brief exposure or static photographs—particularly experiments on the recognition of emotions. Tagiuri (1969) points out that static photograph methodology has limitations, at least in this field, when he says: 'it is rare that judgement of an emotion ever takes place in the same way as it does when one examines a photograph of a facial expression snapped at 20 milliseconds. Historically speaking, we may have been done a disservice by the pioneering efforts of those who, like Darwin, took the human face in a state of arrested animation as an appropriate stimulus situation for studying how well we recognize human emotions' (p. 404). Frijda's (1953) laboratory study supports Tagiuri's conclusion. Judgements of filmed expressions were compared to 10 sec. exposures of still photographs representing the culmination point of the expressions. An increase in accuracy of nearly 50 per cent was found in favour of the films.

Clearly film is better and this is probably so because it is more realistic. The whole series of experiments on the perception of personality and the perception of emotion from still photographs reduce stimuli to a point at which they misrepresent reality.

Person perception experiments like these suggest that subjects rely heavily on certain physical cues when making judgements. In a more realistic situation these cues are probably so minor that they contribute little to the perceptual judgements made.

We are indebted to Florisse Alkema for help with the statistics, students from St Clare's Hall for acting as target persons, and to the SSRC for financial support.

REFERENCES

FRIJDA, N. H. (1953). The understanding of facial expression of emotion. *Acta psychol.* **9**, 294–362.

SCODEL, A. & AUSTRIN, H. (1957). The perception of Jewish photographs by non-Jews and Jews. *J. abnorm. soc. Psychol.* **54**, 278–280.

TAGIURI, R. (1969). Person perception. In G. Lindzey & E. Aronson (eds.), *Handbook of Social Psychology*. Reading, Mass.: Addison-Wesley.

THORNTON, G. R. (1943). The effect upon judgements of personality traits of varying a single factor in a photograph. *J. soc. Psychol.* **18**, 127–148.

THORNTON, G. R. (1944). The effect of wearing glasses upon judgements of personality traits of persons seen briefly. *J. appl. Psychol.* **28**, 203–207.

Manuscript received 5 *February* 1970

Revised manuscript received 21 *May* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 30–41
Printed in Great Britain

Information Processing in Schizophrenia: The Effect of Varying the Rate of Presentation and Introducing Interference

By D. V. HAWKS

Institute of Psychiatry, London

and K. N. ROBINSON

Mental Health Services, Perth, Western Australia

Payne and McGhie offer explanations of schizophrenic thought disorder which emphasize a breakdown in the selectiveness of Broadbent's filter. Yates has observed that an explanation emphasizing the slowness of information processing in schizophrenia equally well follows from Broadbent's model and accords with the experimental literature.

Eighteen male chronic schizophrenics and nine male psychiatric nurses matched for age and intelligence were asked under one experimental condition (the complete recall condition) to reproduce digits presented dichotically at three different rates of presentation, and under another condition (the interference condition) to reproduce only those digits received through a designated channel. There were four different length lists and four presentations of each list length. The chronic schizophrenics were divided into two groups of nine in accordance with their scores on the Venables–O'Connor Paranoid Scale.

An analysis of variance was performed in the complete recall condition using difference scores obtained by subtracting the number of digits correctly reproduced in order in the second half list from the number reproduced from the first half list. In the interference condition the scores employed were the total number of relevant digits correct, the number of relevant lists completely correct and the number of digits reproduced from the irrelevant channel. While group was a significant source of variance in the complete recall condition, contrary to prediction from Yates's theory, the interaction of group and rate was insignificant. This was the case even when the analysis was repeated using difference scores adjusted for variation in the reproduction of the first half list and only the normal and non-paranoid schizophrenic groups compared. There was some evidence that the recall of the second half list improved for all groups at faster rates of presentation. Nor were the three groups differentially affected by variations in the rate of presentation under the interference condition. As predicted from Payne and McGhie, however, the schizophrenics were most affected by distraction from the irrelevant channel.

It is suggested that at slower rates of presentation the digits presented first in the complete recall condition are rehearsed with the result that the reception of those presented later is confused. The need to consider the total length of time over which information is to be processed is discussed. The view is put that retardation in schizophrenia may constitute a secondary adjustment to a state of information overload and an empirical test of this hypothesis proposed.

Payne (Payne *et al.*, 1959; Payne, 1960) and McGhie (McGhie & Chapman, 1961; Chapman & McGhie, 1962, 1963; McGhie *et al.*, 1965*a,b;* McGhie, 1965) have proposed theories of schizophrenic thought disorder which suggest that the thought-disordered schizophrenic's extreme susceptibility to irrelevant stimulation may be due to the breakdown of a hypothetical filter, the normal operation of which is to screen the stimulation impinging on the organism and admit only that which is relevant (Broadbent, 1958).

In a series of investigations involving both auditory and visual channels Chapman and McGhie have demonstrated that introducing an irrelevant source of stimulation, particularly auditory stimulation, results in schizophrenics displaying a greater distraction effect than normals or non-schizophrenic psychiatric patients. That normal subjects are able to largely ignore irrelevant stimulation was shown by Moray (1959) who found that if subjects were told to attend to (i.e. to shadow) a message received in one channel they remained almost totally unaware of the input to the other channel. This selective block was only overcome if the subject's name was included in the 'irrelevant' message. Payne *et al.* (1970) found that schizophrenics given the same task showed more interference from the irrelevant channel than did non-schizophrenic patients or normals.

Yates (1966*a*, *b*) has pointed out that Broadbent's filter theory offers alternative explanations of schizophrenic thought disorder. In Broadbent's model the primary processing channel has a limited capacity, and input exceeding this limit is diverted into a short-term memory system. The amount of information held in the short-term memory is a function of the rate at which information is passed through the primary processing channel. Yates has observed that if the rate at which schizophrenics process information is abnormally slow the result is a greater diversion to the short-term memory system, which by definition can only hold information for a short period of time. As a consequence there is a greater loss of input than occurs in the case of normals. Only part of the relevant stimulation impinging on the individual is successfully processed and the resultant fragmented output appears disordered and bizarre.

The finding of Payne *et al.* (1970) that the rate of presentation had to be considerably slowed before schizophrenics were able to successfully shadow a message even without interference and Chapman & McGhie's (1963) observation that thought disorder in schizophrenics is minimized if the rate of input is slowed is consistent with Yates's hypothesis.

Court & Garwoli (1968) claim to have provided a test of Yates's hypothesis in examining schizophrenic reaction time with increasing levels of complexity. While they demonstrated that the trends describing the reaction times of normals and schizophrenics did not diverge significantly with increasing numbers of choices (there was a significant quantitative difference), in that they did not vary the rate at which information was presented their investigation has only indirect bearing. Karras (1967) found that schizophrenics were slower than controls (non-schizophrenic psychiatric patients) on a simple and two choice reaction-time task. That the chronic schizophrenics were not differentially slowed on the two choice task is interpreted by Karras as indicating that the slower performance of schizophrenics cannot be attributed to their processing data more slowly than non-schizophrenics. Court (1967) has commented on the inadequacy of Karras's design and the consequent ambiguity of his results.

The fact that McGhie and Yates specifically apply their different explanations to non-paranoid schizophrenics makes them directly comparable. In so far as Yates holds that schizophrenics' filter mechanism is unimpaired their recall of information which is presented dichotically should not be affected by the introduction of a message which the subject is instructed to ignore. On the other hand increasing the rate at which the relevant input is presented should result in a significant deterioration in the reproduction of schizophrenics' since Yates holds that the schizophrenic is abnormally slow in processing information through the primary channel. While not contradicting Yates's hypothesis Chapman and McGhie argue that it is principally the schizophrenics' inability to exclude irrelevant input which results in their impaired performance, an impairment which is only lessened to some extent by slowing the rate of presentation.

METHOD

The dichotic stimulation technique initially employed by Broadbent provides a direct test of the two models in that it involves both the primary processing channel and the short-term memory system. Presenting digits dichotically at a rate of at least one pair per second favours ear by ear recall. As a test of Yates's theory requires that some of the input be diverted to the short-term memory system, a form of presentation in which input from one ear (half list) is held in store while the other is processed through the primary processing channel is indicated. If different but simultaneous input is presented to the two ears, and the subject required to recall only one, it is possible to assess the effects of interference at different rates of presentation and the interaction of these two effects.

All subjects were given a standard digit span forward test presented dichotically and increasing in length from one to eight digits with two presentations of each length followed by two practice presentations in which a digit was presented to the left ear succeeded by a different digit to the right ear; the order of presentation being varied. When these digits had been correctly recalled they were presented simultaneously and repeated if necessary until a correct response was made. The test itself comprised two conditions, presented in the same order, the morning and afternoon of the same day.

Complete recall condition

In the complete recall condition 48 lists of digits of four different lengths, two, four, six and eight digits (i.e. 1, 2, 3 and 4 digits per half list or ear), were presented dichotically at three different rates of presentation. The three rates were one pair of digits per second, one pair per half second and one pair per one third second.[*] A rate of one pair per second is the slowest rate at which ear by ear reproduction is strongly favoured. There were four presentations of each length list. Under the complete recall conditions subjects were asked to recall all digits presented.

Lists were chosen from tables of random numbers with the proviso that no digit was included more than once in any list. The order of presentation was the same for all subjects beginning with the shortest length and the slowest rate of presentation. This standard order of presentation would be expected to reduce variations in performance due to differences in list length and rate of presentation since it provided for a gradual and continuous accommodation to the increasing difficulty of the task. Four different length lists were employed as it has been found that the longer the list the more pronounced the serial order effect occurring in ear by ear reproduction and hence the greater involvement of the short-term memory system (Inglis & Sanderson, 1961; Inglis & Caird, 1963).

Interference condition

The interference condition differed from the complete recall condition in that subjects were instructed to recall only those digits presented to a designated ear. For each rate and length combination the relevant ear was specified for each subject by the experimenter and the subject asked to acknowledge this by touching the appropriate ear. The particular rate, length and ear combination employed in each instance was determined in such a way as to ensure that each subject experienced different combinations and all possible combinations were equally represented. In this way it was hoped to control possible differences in lateral dominance. A volume balance was achieved during the practice trials.

---

[*] Rate of presentation refers to the precise timing of simultaneous presentations of pairs of digits. For economy of expression, however, further reference to particular rates will be made in the less specific form of 1, 2 or 3 pairs of digits per second.

Subjects. The subjects were eighteen male schizophrenics selected from three wards of a large psychiatric hospital and nine male psychiatric nurses drawn from the same hospital and restricted age range (23–50 years). All subjects included in the study had IQs of at least 90 on the Mill Hill Vocabulary Scale (Raven, 1948). The schizophrenics had all been continuously hospitalized for at least two years prior to testing. Each patient considered for inclusion was reinterviewed by their doctor and their diagnosis confirmed before the patient was included in the sample. Any patient whose diagnosis included reference to organicity, alcoholism or mental deficiency was excluded. Similarly any patient who had been given ECT in the three months prior to testing, or who had over a longer period been given extensive ECT treatment, was excluded.

Because of the evidence suggesting that paranoid and non-paranoid schizophrenics differ in their manifestation of overinclusive thought disorder (Payne & Hewlett, 1960; Payne *et al.*, 1964) and Yates's assertion that his theory is likely to be more applicable to non-paranoid schizophrenics (Yates, 1966*a*) the patients were rated by the charge nurse on their ward on a scale of paranoid schizophrenia (Venables & O'Connor, 1959). This scale consists of four items found to reliably differentiate paranoid and non-paranoid schizophrenics. The patients were rated toward the end of a ten-day period off medication.

The ratings given by the charge nurses were found to have reasonable internal consistency as measured by inter-item correlations (0·35–0·84). The ratings obtained by each patient on the four items were summed to give a total score and the total scores of all subjects ranked. The nine subjects scoring lowest comprised the non-paranoid group and the nine scoring highest made up the paranoid group.

Table 1. *Comparison of the three groups in terms of age, years of hospitalization, ratings on the paranoid scale and digit span*

	Normals	'Paranoid' schizophrenics	'Non-paranoid' schizophrenics
Age (years)			
Mean	37·88	35·89	39·89
S.D.	7·34	10·05	6·35
Range	24–48	23–50	26–47
n	9	9	9
Continuous hospitalization (years)			
Mean		6·44	9·11
S.D.		3·10	8·11
Range		2–12	2–30
n		9	9
Total ratings on Venables–O'Connor Scale			
Mean		11·33	4·78
S.D.		4·35	0·79
Range		7–18	4–6
n		9	9
Digit span (raw score)			
Mean	6·78	5·89	5·89
S.D.	1·03	0·87	0·99
Range	5–8	4–7	4–8
n	9	9	9

Uncorrelated *t*-tests calculated between the three groups showed that there were no significant differences as regards age or digit span performance. The two schizophrenic subgroups did not differ as regards chronicity ($t = 0{\cdot}87$) but were significantly differentiated on the Venables–O'Connor scale ($t = 4{\cdot}19$; $P < 0{\cdot}01$).

All subjects included in the investigation were tested by a qualified audiometrist in order to ensure a minimum hearing level, defined as within normal limits for the frequencies employed in the investigation. Approximate equivalence between the ears was ensured.

Scoring

Moray & Barnett (1965) have demonstrated that much of the controversy over findings obtained by investigators using dichotic stimulation results from the different methods of scoring employed. They point out that correlations calculated between different scores derived even from the same experiment vary considerably with the consequence that results obtained from different experiments cannot be readily compared. For this reason a number of different scores were employed in the present instance.

Complete recall condition. The score which is most pertinent to Yates's hypothesis is the number of digits reproduced in correct order per half list (i.e. per ear in a dichotic presentation). As it is the second half list in an ear by ear reproduction which is temporarily held in the short-term memory system Yates's theory would hold that with increasing rates of presentation the accuracy of recall of the second-half list should deteriorate in schizophrenics to a greater extent than normals. Reproduction of the first half list would not be expected to deteriorate except at very fast rates of presentation or when longer lists of digits are employed.

Interference condition. Three scores were obtained in the interference condition: (i) the total number of relevant digits correctly reproduced taking order into account ('digits correct'), and (ii) the number of relevant lists correctly reproduced ('lists correct').

In so far as McGhie and Chapman hold that schizophrenics (particularly non-paranoid schizophrenics) are unable to exclude irrelevant input it should follow that the schizophrenic group will correctly reproduce fewer of the relevant digits under the interference condition when order is taken into account. Yates's model on the other hand would lead one to predict significant differences between the groups as a consequence of their differing ability to process information presented at different rates. That is, there should be a significant interaction between group and rate, with the non-paranoid schizophrenics performing worse than the paranoid schizophrenics and normals as a result of their inability to process information quickly.

Both these scores were derived by summing the scores obtained for the two longest length lists (i.e. the 3 and 4 digit lists) thus doubling the number of observations on which each score is based. This resulted in the elimination of length as a source of variance in the analysis of the interference data.

(iii) A third score was obtained by counting the number of digits reproduced from the irrelevant ear ('digits intruded'). A single score for each subject at each of the three rates of presentation was obtained by summing the score got from all four length lists.

In so far as Yates's hypothesis holds that schizophrenics do not suffer any impairment of their selective attention, his model would not predict significant differences between the groups on this score. McGhie and Chapman's model on the other hand would predict that non-paranoid schizophrenics in particular and schizophrenics in general will include more of the digits from the irrelevant ear in their reproductions.

RESULTS

Complete recall condition

In the complete recall condition ear by ear reproduction was favoured and this order of reproduction was taken to be the correct or required order. The first digit reproduced was taken to indicate which of the two half lists had been given preference by the subject. On those occasions when the first digit reproduced was not the first digit of either half list (this occurred in 3·47 per cent of 1296 lists) both combinations were compared with the list as reproduced and a score given which reflected the best match.

Table 2. *Mean digits correctly reproduced from the first and second half lists at the three rates of presentation*

		One pair per second			Two pairs per second			Three pairs per second		
No. of digits in list	Half list reproduced	Normals	Paranoid schizo-phrenics	Non-paranoid schizo-phrenics	Normals	Paranoid schizo-phrenics	Non-paranoid schizo-phrenics	Normals	Paranoid schizo-phrenics	Non-paranoid schizo-phrenics
2	First	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00
	Second	1·00	1·00	0·92	1·00	1·00	0·97	1·00	0·97	0·97
4	First	1·53	1·61	1·42	1·92	1·81	1·56	1·92	1·92	1·78
	Second	1·31	1·17	0·97	1·86	1·75	1·31	1·89	1·81	1·28
6	First	1·92	2·03	1·67	2·78	2·47	2·03	2·78	2·53	2·25
	Second	1·36	0·69	0·58	2·28	1·39	0·81	2·17	1·08	0·92
8	First	2·47	2·11	1·86	3·36	2·83	2·36	3·53	2·89	2·83
	Second	0·89	0·36	0·31	1·22	1·00	0·33	1·83	0·92	0·56

It would be predicted from Yates's theory that rate of presentation and diagnosis would prove significant sources of variance; specifically Yates's hypothesis would predict a significant interaction between these effects, with schizophrenics showing increasing inaccuracy of recall of the second half list reproduced at faster rates of presentation. Further, the tendency for schizophrenics to become increasingly inaccurate at faster rates of presentation should be most apparent in those patients having lower scores on the paranoid scale.

The fact that the number of digits correctly recalled is a function of the length of the list makes interpretation of the length effect and its interactions impossible except that the scores are expressed at some percentage of the maximum possible score. Since such a transformation is confounded by the fact that the longer the list the greater the number of digits employed and hence the fewer errors of commission possible the data were not transformed and the length effect was left uninterpretable in the analysis.

In order to test predictions from Yates's theory an analysis of variance was performed on the scores obtained by subtracting the number of digits correctly reproduced (in order) from the second half list (X_2) from the comparable score obtained on the first half list (X_1).

Table 3. *Analysis of variance of difference scores (number of digits correctly reproduced in the first half list—number of digits correctly reproduced from the second half list) X_1–X_2, unadjusted for differences in X_1: complete recall*

Source	D.F.	*F*	*P*
Group (G)	2	6·59	< 0·05
Rates (R)	2	0·40	n.s.
Length (L)	3	149·87	< 0·01
R × L	6	1·23	n.s.
G × R	4	0·65	n.s.
G × L	6	2·26	< 0·05
G × L × R	12	0·26	n.s.
Subjects within groups	288		
Total	323		

While group proved a significant source of variance neither the rate of presentation nor the interaction of group × rate were statistically significant effects. Consideration of the mean difference scores obtained by the three groups, however, suggests that the non-paranoid schizophrenics tended to perform worse at faster rates of presentation.

Table 4. *Mean difference scores obtained by the normal and schizophrenic groups at the three rates of presentation*

	Rate		
	1	2	3
Normals	2·36	2·69	2·33
Non-paranoid schizophrenics	3·30	3·53	4·11
Paranoid schizophrenics	3·50	3·03	3·55

Accordingly an analysis of covariance was performed using the normal and non-paranoid schizophrenic groups only as Yates's hypothesis relates specifically to non-paranoid schizophrenics. In order to exclude the possibility that differences in X_1–X_2 merely reflect differential performance on the first half list, this analysis was performed on difference scores adjusted for variation in X_1.

Table 5. *Analysis of covariance of difference scores (X_1–X_2) adjusted for variation in X_1: normal and non-paranoid schizophrenic group only*

Source	D.F.	F	P
Group (G)	1	26·199	< ·001
Rate (R)	2	0·304	n.s.
Length (L)	3	34·250	< ·001
G × R	2	0·646	n.s.
G × L	3	3·477	< .05
R × L	6	0·306	n.s.
G × R × L	6	0·288	n.s.
Within cells	191		
Total	214		

While diagnosis and length proved significant sources of variance when adjusted scores were employed, the interaction between rate and diagnosis did not attain statistical significance. It may be concluded therefore that the groups did not differ in their reproduction of the second half list as a function of the rate of presentation as Yates's theory would require, though it might be argued that the small numbers involved when only the normal and non-paranoid groups were compared mitigates against finding significant results.

Interference condition

Relevant digits correct. Comparing the groups on the number of digits correctly reproduced from the relevant ear showed that whereas there were significant differences between the groups neither the rate of presentation nor the interaction of rate and group were significant sources of variance.

Table 6. *Analysis of variance of the number of 'relevant' digits correctly reproduced (order taken into account): interference condition*

Source of variance	D.F.	F ratio	P
Group	2	20·001	< 0·005
Rate	2	0·097	> 0·25
Group × rate	4	0·204	> 0·25
Within cells	72		
Total	80		

Applying t-tests to the differences between the means of the groups demonstrates that normals obtained significantly higher scores than either the paranoid or non-paranoid schizophrenic groups ($t=3\cdot889$; $P<0\cdot01$; $t=10\cdot812$; $P<0\cdot01$ respectively). The paranoid schizophrenics for their part obtained higher scores than the non-paranoid schizophrenics ($t = 6·924$; $P < 0·01$).

Relevant lists correct. As would be expected similar results were obtained when the groups were compared in terms of the number of relevant lists correctly reproduced. Differences between the three groups were assessed for significance using t-tests. Normals obtained significantly higher scores than either of the two schizophrenic subgroups ($t = 6·007$; $P < 0·01$; $t = 11·206$; $P < 0·01$ for the paranoid and non-paranoid subgroups respectively.) The paranoid schizophrenics obtained significantly higher scores than their non-paranoid counterparts ($t = 5·199$; $P < 0·01$).

Irrelevant digits intruded. The score most pertinent to a test of Chapman and McGhie's hypothesis is the number of digits reproduced from the irrelevant ear. If, as Chapman and McGhie contend, schizophrenics are unable to exclude irrelevant stimuli their reproductions should contain more of the digits presented to the irrelevant ear.

Table 7. *Analysis of variance of the number of 'irrelevant' digits reproduced*

Source of variance	D.F.	F ratio	P
Group	2	63·055	< 0·005
Rate	2	0·196	> 0·25
Group × rate	4	0·283	> 0·25
Within cells	72		
Total	80		

The three groups did in fact differ markedly in the number of digits reproduced from the irrelevant ear; on the other hand rate of presentation and the interaction between rate and group proved insignificant sources of variation. t-tests applied to the differences between the mean scores obtained by the three groups showed that the normal group reproduced significantly fewer irrelevant digits than either the paranoid or non-paranoid group ($t = 3·661$; $P < 0·01$; $t = 3·003$; $P < 0·01$ respectively). While on average the non-paranoid schizophrenics reproduced twice as many irrelevant digits as the paranoid subgroup the wide distribution of the non-paranoid scores mitigated against finding a significant difference between the two subgroups ($t = 1·64$; $P > 0·1$).

It could be argued that the results obtained using this score tell us nothing about the subjects' attention to the irrelevant ear since with only a limited number of digits available almost any error of reproduction will count as an interference error. Thus in the three digit list there are only three digits, which, if reproduced, would not count as interference errors; in the four digit list the possibility is reduced to one. As a consequence all the results may show is that schizophrenics, particularly non-paranoid schizophrenics, make more errors of reproduction than normals and that in the present circumstances some of these errors will be labelled interference errors.

Repeating this analysis for only the one and two digit lists combined, however (where the number of potential non-interference errors is greater and the number of errors properly regarded as interference errors smaller), substantiates the results obtained using the three and four digit lists. This and the fact that, as predicted from Chapman and McGhie, non-paranoid schizophrenics make the most interference errors suggest that it is an inability to exclude irrelevant input which results in the schizophrenics' poorer performance in the present task. Even so the limits placed upon any test of the interference theory when digits are employed and varying length lists used is a consideration which should not be overlooked.

DISCUSSION AND CONCLUSIONS

The results of the present investigation provide substantial support for Chapman and McGhie's theory. Schizophrenics performed significantly worse than normals in a situation requiring attention to one of two auditory messages presented dichotically. As predicted the non-paranoid schizophrenics tended to do worst in these circumstances. A direct measure of attention to the irrelevant ear (the number of digits reproduced from the irrelevant ear) showed that schizophrenics included significantly more irrelevant digits in their reproductions.

By contrast under neither condition were the three groups differentially affected by variations in the rate of presentation, as Yates's theory would predict. There was no evidence that schizophrenics made more errors when reproducing the second half list at the three pairs per second rate of presentation than they made at the one pair per second rate. There was in fact some tendency for the recall of the second half list to improve at faster rates of presentation for all three groups.

It is conceivable that at slower rates of presentation those digits presented first are rehearsed with the result that the reception of those presented later is adversely affected in the complete recall situation. To the extent that a faster rate of presentation precludes rehearsal there may be less interference with incoming information at these rates.

While the effect of list length on reproduction has not been considered in the present analysis it is of obvious relevance to Yates's hypothesis. Reproducing six digits presented at the rate of one per second is a less difficult task than repeating eight digits presented at the same rate if only because the same rate of processing must be sustained longer in the latter case. Insofar as the schizophrenics' slowness in processing information results in a build-up of input in the short-term memory system the decay of information from the short-term memory system should increase with longer series of digits even though the rate of presentation remains constant. The fact that schizophrenics give irrelevant responses even when the rate of presentation is arbitrarily slowed suggests that the total period of time over which information is presented is an important variable.

The fact that neither of the schizophrenic sub-groups reproduced more of the irrelevant digits in the interference condition at faster rates of presentation contradicts Chapman & McGhie's (1963) observation that if the rate of presentation is slowed schizophrenics show less impairment of their selective attention. Increasing the rate of presentation should exacerbate schizophrenics' attention defect. Even so the suggestion that schizophrenics' attention defect can be minimized by slowing the rate at which information is presented is worthy of further examination. Payne & Hewlett (1960) found that acute schizophrenics who were non-overinclusive on their tests tended to be retarded. It is possible that the retardation observed in some schizophrenics is an adjustment to a condition of information overload. If the rate at which information is received affects the ability of the schizophrenic to screen out irrelevant stimuli it should be possible to manipulate the degree of interference manifest by schizophrenics by varying the rate at which they are required to process information. A test of such a proposition would be to require retarded schizophrenics to process information at a rate in excess of their chosen rate and conversely to experimentally retard the rate at which overinclusive schizophrenics process information. It would be predicted that retarded schizophrenics would show increased effects of interference and overinclusive schizophrenics manifest less interference.

K. N. Robinson was in receipt of a Western Australian Government Traineeship during the course of this study. The facilities used were made available by the Western Australian Mental Health Services and the University of Western Australia. Miss Beryl Skinner provided secretarial assistance.

REFERENCES

BROADBENT, D. E. (1958). *Perception and Communication*. London: Pergamon Press.

CHAPMAN, J. & McGHIE, A. (1962). A comparative study of disordered attention in schizophrenia. *J. ment. Sci.* **108**, 487–499.

CHAPMAN, J. & McGHIE, A. (1963). An approach to the psychotherapy of cognitive dysfunction. *Br. J. med. Psychol.* **36**, 253–260.

COURT, J. H. (1967). Comments on Karras. *Psychon. Sci.* **8**, 548.

COURT, J. H. & GARWOLI, E. (1968). Schizophrenic performance on a reaction time task with increasing levels of complexity. *Br. J. soc. clin. Psychol.* **7**, 216–223.

INGLIS, J. & CAIRD, W. K. (1963). Age differences in successive responses to simultaneous stimulation. *Can. J. Psychol.* **17**, 98–105.

INGLIS, J. & SANDERSON, R. E. (1961). Successive responses to simultaneous stimulation in elderly patients with memory disorder. *J. abnorm. soc. Psychol.* **62**, 709–712.

KARRAS, A. (1967). The effect of stimulus-response complexity on the reaction time of schizophrenics. *Psychon. Sci.* **7**, 75–76.

McGHIE, A. (1965). Psychological studies of schizophrenia. In T. Freeman, J. L. Cameron & A. McGhie. *Studies on Psychosis*. London: Tavistock Publications.

McGHIE, A. & CHAPMAN, J. (1961). Disorders of attention and perception in early schizophrenia. *Br. J. med. Psychol.* **34**, 103–116.

McGHIE, A., CHAPMAN, J. & LAWSON, J. S. (1965*a*). The effect of distraction on schizophrenic performance. 1. Perception and immediate memory. *Br. J. Psychiat.* **111**, 383–390.

McGHIE, A., CHAPMAN, J. & LAWSON, J. S. (1965*b*). The effect of distraction on schizophrenic performance. 2. Psychomotor ability. *Br. J. Psychiat.* **111**, 391–398.

MORAY, N. (1959). Attention in dichotic listening: affective cues and the influence of instructions. *Q. J. exp. Psychol.* **11**, 56–60.

MORAY, N. & BARNETT, T. (1965). Stimulus presentation and methods of scoring in short-term memory experiments. *Acta psychol* **24**, 253–263.

PAYNE, R. W. (1960). Cognitive abnormalities. In H. J. Eysenck (ed.), *Handbook of Abnormal Psychology*. London: Pitman.

PAYNE, R. W., CAIRD, W. K. & LAVERTY, S. G. (1964). Overinclusive thinking and delusions in schizophrenic patients. *J. abnorm. Psychol.* **68**, 562–566.

PAYNE, R. W. & HEWLETT, J. H. G. (1960). Thought disorder in psychotic patients. In H. J. Eysenck (ed.), *Experiments in Personality*, vol. 2. London: Routledge & Kegan Paul.

Payne, R. W., Hochberg, A. C. & Hawks, D. V. (1970). Dichotic stimulation as a method of assessing disorder of attention in overinclusive schizophrenic patients. *J. abnorm. Psychol.* **76**, 185–193.

Payne, R. W., Matussek, P. & George, E. I. (1959). An experimental study of schizophrenic thought disorder. *J. ment. Sci.* **105**, 627–652.

Raven, J. F. (1948). *The Mill Hill Vocabulary Scale.* London: Lewis.

Venables, P. H. & O'Connor, N. (1959). A short scale for rating paranoid schizophrenia. *J. ment. Sci.* **105**, 815–818.

Yates, A. J. (1966*a*). Data processing levels in thought disorder in schizophrenia. *Aust. J. Psychol.* **18**, 103–117.

Yates, A. J. (1966*b*). Psychological deficit. *Ann. Rev. Psychol.* **17**, 111–144.

Manuscript received 9 June 1969

Revised manuscript received 27 July 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 42–54
With 1 *text-figure*
Printed in Great Britain

The Eysenck Personality Inventory and Social Desirability Response Set with Student and Clinical Groups

BY E. E. RUMP AND JOHN COURT

University of Adelaide

The EPI was administered to three student groups and to two clinical groups in studies allowing the investigation of Social Desirability (D) effects. Form A of the EPI and the Edwards D scale were completed by 100 students; Form A and the Crowne–Marlowe D scale by 177 students; Form B and the Crowne–Marlowe D scale by 175 students; and Form A, including the Lie (L) scale, by 38 potential renal transplant patients and 15 potential renal donors. Negative linear correlations between the Neuroticism (N) scale and the D or L scales were found for all groups. A very pronounced reduction in N scores was found with potential renal donors, together with high L scores. In most cases, no reliable correlation between Extraversion (E) and the D or L scales was found. The N and E scales were correlated negatively for Form B. Implications of these results for the use of the EPI are considered. Correlations with age and IQ were also obtained.

It has been stated that '"Desirability response set" . . . has not been found to play an important part under ordinary conditions' with regard to the EPI N and E scales (Eysenck & Eysenck, 1964). However, a number of reports have shown that various personality scales are affected by Social Desirability response set (D), following the work of Edwards (1957). In some studies a Lie scale (L) has been used as a measure of D responding, while in others special D scales have been devised with items having a much greater scoring-response probability than L scale items. The MPI has been investigated in this respect, and since the EPI was developed from it and gives similar results (Bailey & Metcalfe, 1969), the findings are relevant. Gibson (1962) obtained significant correlations of −0·29 and −0·21 between the full MPI N scale and an L scale for English and U.S.A. subjects respectively. Eysenck & Eysenck (1963*a*) used short N and L scales, and found the regression was linear and highly significant with normal instructions.

Despite this evidence, Eysenck & Eysenck concluded: 'For the great majority, desirability response set may be disregarded.' Three justifications for this conclusion were given: the L–N relationship was slight for low values on the L scale; the mean L score was low with normal instructions; and a large increase in L and decrease in N scores occurred with instructions to 'fake good'. However, this fakeability of N scores might be considered to *confirm* the possibility of socially desirable responding; and the low mean L score may well indicate a scale with little discrimination at the low end, rather than showing the unimportance of the D variable. Over half of the subjects scored 1 or 0 on the L scale, and such poor discrimination arises inevitably from the deliberate selection of items giving a very low probability of an L-scoring response.

With groups having stronger motivation to 'look good', higher mean L scores have been obtained, and quite clear relationships between N and L scales have then been found. Thus apprentices tested at work gave a significant correlation of −0·36 in Gibson's study; male psychopaths gave a significant correlation of −0·31 between EPI (Form A) L and N scales (McKerracher & Watson, 1968); depressed patients gave a significant correlation also (Bailey & Metcalfe, 1969); male cancer patients showed a significant association between the L scale and a 10-item N scale (Huggan, 1968); and a marked decrease in N (Form A) with increasing L scores was found with anxious and neurotic patients (Knowles & Kreitman, 1965). Moreover, Williams (1969) showed that the rank correlation between mean N and L scores across eight varied groups was as high as −0·79.

However, when studying normal groups with low L scores, it is important that the social desirability effect should be tested by means of a D scale having adequate power of discrimination. This may be achieved by having more items in total, and a smaller proportion of extremely improbable items. Feather (1967) used the Crowne & Marlowe (1960) scale, and found a significant correlation of −0·68, with the MPI N scale, for female students only. Farley (1966) investigated the EPI (Form A), using both the Crowne & Marlowe and the Edwards D scales: correlations with N of −0·30 and −0·60 respectively were found.

This accumulation of evidence is fairly consistent, but further study of the EPI is required with normal groups tested under conditions designed to minimize deliberate 'faking'. On the other hand, the possible extent of the depression of N scores due to social desirability needs also to be determined under conditions of very high motivation to 'look good' but with normal instructions. These were the two main aims of the investigations to be described.

Two additional topics for investigation were the effects of social desirability on the EPI E scale, and the relationship between the E and N scales. There is some uncertainty on both these matters. Eysenck & Eysenck (1963*a*) found a significant linear reduction in E scores with increasing L scores, although they also suggested that 'faking' leads to *less extreme* E scores whether high or low. Farley (1966) similarly found a negative correlation between the EPI (Form A) E scale and the Crowne & Marlowe D scale, but Gibson (1962) found non-significant correlations between the MPI E scale and an L scale. Knowles & Kreitman (1965) also found a non-significant trend, using the EPI (Form A) E and L scales with neurotic patients. Bailey & Metcalfe (1969), too, reported no relationship between L and E scales, and yet McKerracher & Watson (1968) found a correlation of −0·50 for female psychopaths. However, two *positive* correlations have been reported. Blackburn (1965) found that a 'repression of symptoms' scale, derived from the MMPI, correlated +0·33 with the MPI E scale for neurotic patients, and Feather (1967) obtained a correlation of +0·32 between the Crowne–Marlowe D scale and the MPI E scale with female students.

This diversity in the results for E scales could be due to the 'dual nature of extraversion' (Eysenck & Eysenck, 1963*b*; Sparrow & Ross, 1964); or to the social desirability effect being bidirectional in the avoidance of both extremes (Eysenck & Eysenck, 1963*a*); or to the direction and extent of the effect being influenced by some modifier variable such as social context (McKerracher & Watson, 1968). An attempt was therefore made specifically to test the first two hypotheses by scoring appropriate subscales from EPI E items for the two largest groups tested in the present study.

The relationship between EPI E and N scales is also of interest. The manual suggests that the correlations between E and N scales 'are near enough to zero to make the scales for all practical purposes independent'. This is correct for Form A, on the basis of the correlation given for 2000 normal subjects; but for Form B, the correlation is −0·116, which has a two-tailed probability of less than 10^{-6}. With conventional levels of confidence, the hypothesis of orthogonality for EPI (Form B) N and E scales may therefore be rejected. The Form B correlations for neurotics and psychotics reported in the manual are also negative and significant at the 5 per cent level. Correlations of −0·26, −0·32 and −0·33 between N and E (Form B) scales were found by Richardson (1968) in three groups of army recruits. Smaller negative correlations were also found for Form A scales with these groups. Farley (1967), Bailey & Metcalfe (1969) and Colston (1969) reported no significant correlation between EPI (Form A) E and N scales, although McKerracher & Watson (1968) obtained a correlation as high as −0·58 with female psychopaths. Thus it would appear that the correlation is often negative, but with normal groups is marked only with Form B. Further evidence on this issue, and on the social desirability effect, was sought in the investigations described below.

METHOD

The EPI was administered to three student groups and to two clinical groups. Group 1 was a sample of University of Adelaide students enrolled for first-year psychology in 1967. The sample was a random one, except that only those less than 30 years old and with English as their native language were invited to participate; of the 111 eligible subjects who were invited, 11 declined owing mainly to timetable difficulties. The initial letter emphasized that the volunteers should be 'diligent and sincere', that results would be confidential, and that subjects would be given an outline of their own results on request. In this way, a form of 'client contract' was established (Lovell, 1967), so as to minimize deliberate faking. Subjects were tested in six subgroups, and given the EPI (Form A) and the Edwards D scale (labelled 'S.D. Questionnaire'), together with other tests, including the ACER Advanced Test AL as a measure of IQ.

The Edwards scale has been criticized because it confounds D responding *per se* with psychopathological tendencies (Crowne & Marlowe, 1960). Because of this, the Crowne–Marlowe D scale, which was designed to be independent of psychopathology, was used with two additional groups of first-year psychology students. All those attending the first week's tutorial classes in 1969 were tested, except students who were repeating the course or who had previously completed the EPI. It was explained that the experience of answering the questionnaires would be useful in connexion with later lectures, and that names should not be written on the forms so that confidentiality was ensured. The students were asked to 'answer the questionnaires seriously and sincerely, because the research is intended to investigate their usefulness.' Again it was hoped that these instructions would maximize cooperation, and indeed there were no 'spoilt' forms. The first ten classes formed group 2, and were given the EPI (Form A), followed by the Crowne–Marlowe D scale (labelled 'Personal Reaction Inventory'). The other 11 classes formed group 3 and were given the EPI (Form B) followed by the same D scale. There were 177 in group 2 and 175 in group 3. The testing of groups 1, 2 and 3 was personally supervised by the first author, with the aid of the class tutors in the latter groups.

Two clinical groups were tested by the second author. The standard test instructions were given but conditions were maximally conducive to D responding. Of all reasons why a person might wish to 'look good' on a test, none could be more compelling than the belief that the issue was one of life or death. Group 4 consisted of 38 potential patients, and group 5 consisted of 15 potential donors in a renal homotransplantation programme, in which psychological evaluation, including the EPI (Form A), played an integral part. The patients were tested in the course of an extensive series of investigations (Cramond *et al.*, 1968), and were aware that there were far more patients than could be given either recurrent dialysis or transplantation. The psychological tests were among the final investigations, so that patients were highly anxious, but also strongly motivated to present themselves in a favourable light. (In fact, the EPI scores were used to observe group trends and not for individual assessment.)

Potential renal donors, it might be supposed, would not have been so strongly motivated. Yet they found themselves exposed to a six-stage series of screening investigations, including the EPI (Form A), like the potential patients. Donors often expressed surprise at this screening, but it was also evident that, by the time they came to the psychologist at the fifth stage, the 15 people concerned were determined not to 'fall at the last fence'. It may also be noted that while the patients had had much prior contact with members of the hospital team whom they had grown to trust, donors had had little contact prior to the screening procedures.

Scoring

The questionnaires for groups 1, 4 and 5 were scored using stencils. Those for groups 2 and 3 were scored by means of FORTRAN programs written by the first author: data cards were punched direct from the questionnaires after editing for completeness, and run on a CDC 6400 computer. Thirteen of the subjects in groups 2 and 3 had left a few items unanswered. Rather than reject the subjects, the 19 missing responses involved were completed by a random process.

In addition to E, N, L and D scale scores, certain subscale and derived scores were obtained for groups 2 and 3. Because all items on the EPI N scale are keyed for a 'yes' response, and since there are slightly more 'true-scoring' items on the Crowne–Marlowe D scale, a control for Acquiescence response set was desirable. To achieve this, D(T) and D(F) subscales were scored separately, being based on the 18 'true-scoring' items and 15 'false-scoring' items of the Crowne–Marlowe scale, respectively.

To aid the interpretation of any Desirability effect on the E scale, the deviation of each score from the mid-point of the scale was obtained. The mid-point is 12, and this is also approximately the mean score for normal subjects on the scale. Thus $E(\text{dev.}) = (E \sim 12)$.

In addition, subscales E(imp.) and E(soc.) were also derived, to measure separately the impulsivity and sociability components of extraversion which have been demonstrated by factor analysis. It was considered that factor loadings would not be the best method of subdividing items, since face validity was considered more relevant than factorial purity for the present purpose. The distinction between these two criteria is illustrated by the item 'Would you rate yourself as an impulsive individual?', which clearly relates to 'impulsivity' and yet has a zero loading on the factor identified as Impulsivity by Sparrow & Ross (1964). Therefore two judges independently considered each E item, to decide whether degree of 'sociability' or of 'impulsivity' was implicated, or neither. The two judges agreed that the following items refer to sociability: Form A, Qs. 15, 20, 22, 32, 44, 46; Form B, Qs. 5, 10, 15, 29, 32, 37. The following items were considered to refer to impulsivity: Form A, Qs. 1, 3, 5, 8, 10, 13, 27, 34, 39, 41, 49; Form B, Qs. 1, 13, 17, 22, 34, 39, 41, 46, 49, 56. These items formed the subscales E(soc.) and E(imp.), and were keyed in the same way as for the full E scale.

Analyses

Means, $\hat{\sigma}$ values, and the correlation matrices were computed using a CORMAT program. For the student groups, numbers were sufficiently large to allow additional separate analyses for males and females, so that sex differences of relevance could be noted. A QIKPLT program was used to investigate the form of the relationship between N and D scales. Because of the doubt about the relationships with the E scale, the significance of correlations involving E was evaluated by two-tail tests. Relationships between N and D or L scales, and between D and L scales, for which well-founded predictions could be made, were evaluated by one-tailed tests.

RESULTS

Group means

Means and standard deviations are given in Table 1. In general, the data for the student groups were consistent with previous research data, including 'norms' published by the authors of the three tests used. In five out of six cases the mean N score was slightly above the norm for the form used, but this was the case also for students in the standardization sample. Koller & Castanos's (1969) conclusion, that English students were more neurotic than Australian students, was not confirmed.

The mean score for group 1 on the Edwards D scale was 30·4, which is about mid-way between the values of 27·8 and 31·8 obtained for that scale by Edwards and by Crowne & Marlowe respectively, for males and females combined. Similarly the mean scores for groups 2 and 3 on the Crowne–Marlowe D scale were near the mean of 13·72 (with $\sigma = 5{\cdot}78$) reported by those authors.

Females gave a slightly higher mean N score than males (see Table 1). The difference was significant for groups 1 and 3 at the 5 per cent level in each case, and was similar to that reported in the EPI manual. Such a sex difference was reported also for the PEN N scale (Eysenck & Eysenck, 1969).

By comparing the means for groups 2 and 3, Form B was seen to give significantly higher E scores ($P < 0{\cdot}01$ for males and for females), and significantly lower L scores ($P < 0{\cdot}01$ for males and for females), compared with Form A. These differences were consistent with those shown in the manual, and those found by Salas & Richardson (1968) with Australian army applicants.

The means for the clinical groups are of primary interest. It is suggested that the means on the N scale were affected by social desirability responding. Patients and potential donors in a renal transplantation programme often show intense anxiety in psychiatric interviews and projective tests (Cramond *et al.*, 1967). Since N scales reflect anxiety level to some extent (Adcock, 1965), high N scores might be expected from the clinical groups. In fact, the mean N values for male and female patients were about the same as those for the student groups. The donors obtained mean N values very much *lower* than either the 'norm' or the student groups. Because of small numbers, male and female donors were combined to test the significance of these differences: it was found that the differences in mean N (EPI Form A) between the donors and the 'norm', and between donors and the student groups 1 and 2 combined, were significant at the 1 per cent level on two-tailed tests. It is postulated that the donors had such low N scores because of their strong motivation to demonstrate that they were 'fit to serve' and that this motivation led to an increased tendency to respond in the socially desirable direction. The interpretation is supported by the considerably raised L scores for this group above the scale 'norm' ($P < 0{\cdot}01$). The patients also gave high L scores as predicted ($P < 0{\cdot}01$), and it is possible that their N scores also were affected to some extent by desirability responding.

Table 1. Means and standard deviations of EPI and Social Desirability Scales, for student and clinical groups

Group	*Age*		EPI scales*						Social Desirability†	
	Years	$\hat{\sigma}$	E	$\hat{\sigma}$	N	$\hat{\sigma}$	L	$\hat{\sigma}$	D	$\hat{\sigma}$
EPI 'norms'			12·07	4·37	9·06	4·78	2·26	1·57		
				MALES						
Students										
Gp. 1 ($n=50$)	20·88	3·29	12·64	3·60	8·70	3·96	1·92	1·52	31·48	4·66
Gp. 2 ($n=91$)	20·18	5·00	11·32	3·63	10·73	4·98	2·52	1·70	13·40	5·53
Gp. 3 ($n=86$)	19·70	3·35	14·14	4·15	11·21	4·46	1·45	1·36	15·16	5·91
Renal patients										
Gp. 4 ($n=19$)	36·79	11·01	10·37	4·13	10·05	4·16	3·32	2·24		
Renal donors										
Gp.5 ($n=8$)	41·00	13·89	10·38	4·47	3·75	4·20	4·12	1·89		
				FEMALES						
Students										
Gp. 1 ($n=50$)	19·70	1·95	12·24	4·63	10·94	5·03	1·86	1·58	29·36	6·18
Gp. 2 ($n=86$)	19·44	5·01	10·93	3·98	11·92	4·97	2·49	1·67	13·94	5·26
Gp. 3 ($n=89$)	18·36	2·25	13·47	3·74	12·47	3·86	1·37	1·23	14·37	5·43
Renal patients										
Gp. 4 ($n=19$)	33·84	8·61	10·84	4·84	12·00	5·29	2·89	1·85		
Renal donors										
Gp. 5 ($n=7$)	55·86	9·70	11·43	4·24	6·00	3·37	4·14	1·68		

* Form A was used for groups 1, 2, 4 and 5, and for the quoted 'norms'; Form B was used for group 3.

† The Edwards scale was used for group 1; the Crowne–Marlowe scale was used for groups 2 and 3.

Table 2. *Correlations between EPI and Social Desirability Scales for five groups*

Group	Tests used	*r* values for the pairs of scales indicated					
		N & L	N & D	L & D	E & L	E & D	E & N
1 ($n = 100$)	EPI (A), Edwards	−0·25**	−0·79**	0·25**	−0·25*†	0·18	−0·12
2 ($n = 177$)	EPI (A), Crowne-Marlowe	−0·13*	−0·35**	0·60**	−0·13	0·02	−0·10
3 ($n = 175$)	EPI (B), Crowne-Marlowe	−0·08	−0·37**	0·46**	−0·07	0·06	−0·31**
4 ($n = 38$)	EPI (A)	−0·42**			−0·25		−0·19
5 ($n = 15$)	EPI (A)	−0·45*			0·13		−0·02
		One tailed tests			Two-tailed tests		

* $P < 0{\cdot}05$. ** $P < 0{\cdot}01$. † This correlation derives mainly from female subjects for whom $r = -0{\cdot}40$; for males, $r = 0{\cdot}05$.

Correlations between measures of Social Desirability

If L and D scales both measure Social Desirability response set, then they should intercorrelate positively. The correlations were 0·25, 0·60 and 0·46 for groups 1–3 respectively ($P < 0{\cdot}01$ in each case). Possibly the value for group 1 would have been higher but for the restricted range of the L scale, which gave mean values near the zero of the scale, and low standard deviations. This interpretation is supported by the fact that the highest correlation was for group 2, which also had the highest mean L values and the highest standard deviations of the student groups. Therefore, it is suggested, the L scale is a better index of individual differences in D responding with groups having above average mean L values.

Evidence for the internal consistency of the Crowne–Marlowe D scale was provided by the correlation between D(T) and D(F) subscales. The correlation was 0·50 for group 2 and 0·55 for group 3 ($P < 0{\cdot}01$). The consistency is quite good, despite any contrary effect of Acquiescence response set.

Correlations between N and L or D scales

In all three cases there was a strong negative correlation between N and D scales. The correlation between N (Form A) and the Edwards D scale was −0·79; this was

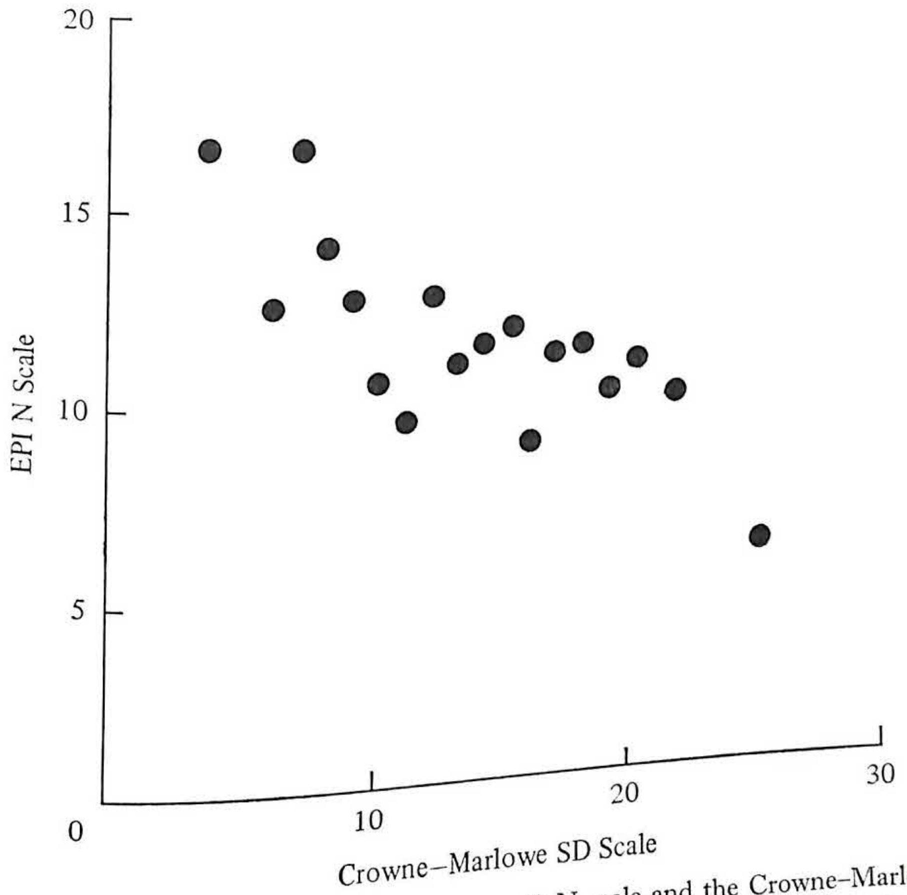


Fig. 1. The relationship between the EPI (Form A) N scale and the Crowne-Marlowe D scale: group 2 (177 subjects). Each point represents the mean for a grouping of at least five subjects.

about as strong as the former's reliability correlation with the N (Form B) scale given in the manual as 0·81 and reported by Salas & Richardson (1968) for Australian subjects as 0·79. It would appear that the Edwards D scale measures 'a maladjustment-psychopathological factor . . . with SD an essential part of it' (Farley, 1968) to about the same extent as the EPI N scale.

On the other hand, the Crowne–Marlowe D scale appears to define a distinct factor of 'approval-seeking' (Farley, 1968). The influence of this factor on the EPI N scale is indicated by its correlation with the Crowne–Marlowe scale. The correlation was −0·35, for Form A and −0·37 for Form B.

The form of the relationship between the D and N scales was inspected. There was an almost linear negative relationship throughout the range, for all three groups. This is illustrated in Fig. 1, which shows the nature of the relationship for the largest group. The linear relationship between social desirability responding and N scales reported by Williams (1969) was therefore supported, rather than the non-linear trend suggested by Gibson (1962).

As pointed out in the previous section, the L scale had rather restricted variance. Nevertheless, there were significant correlations between L and N scales for 4 of the 5 groups, as shown in Table 2. The clinical groups gave L scores with greater variance and less skew than the student groups. Because of this, the L scale was a more satisfactory measure of Desirability responding for the clinical groups, and gave higher correlations (−0·42 and −0·45) with the N scale. On the other hand, group 3 gave a non-significant correlation, and this may have been due to the very restricted range of the L scale with this group: for both sexes combined, mean L was only 1·41, and σ only 1·29.

The D & N correlation and Acquiescence

The effect of Acquiescence response set on the correlation between the Crowne–Marlowe D scale and Neuroticism was determined by considering the D(T) and D(F) subscales. Any negative correlation between the D(F) and N scales would be produced by both Desirability and Acquiescence response sets acting together. On the other hand, any effect of Acquiescence on the D(T) and N correlation would be to decrease the size of the correlation, or even to produce a positive correlation if Acquiescence has a stronger effect than Desirability. Thus if the D(T) and N correlation is less negative than the D(F) and N correlation, this may be taken to indicate the influence of Acquiescence response set.

For group 2 the D(F) and N correlation was −0·39 and the D(T) and N correlation was −0·22. Using Fisher's conversion, and transformation of the difference to a standard-error score, the difference between the two correlations was found to be significant at the 5 per cent level on a one-tailed test. For group 3 the corresponding correlations were −0·48 and −0·20 respectively, and these values were found to be significantly different at the 1 per cent level on a one-tailed test. It is clear that some reduction in the magnitude of the correlation with the N scale occurred using the D(T) subscale, due to Acquiescence response set. Nevertheless, Desirability response set had the stronger effect, since the D(T) and N correlation was still negative ($P < 0{\cdot}01$ in both cases) even with Acquiescence working in the contrary direction.

Correlations between E and L or D scales

There was only one significant correlation between the E scale and any of the Social Desirability scales. This was the correlation of $-0{\cdot}40$ between the EPI (Form A) E and L scales for female students in group 1. However, no consistent correlation was found between E and L scales with other groups, nor was any significant correlation obtained between E and D scales. Moreover, none of the subscales, E(dev.), E(soc.), and E(imp.), was correlated significantly with either D or L scales: none of these correlations was greater than 0·1 for either group 2 or group 3. In view of these results, and evidence from previous research, it may be concluded that there is no general effect of Social Desirability response set on E scales, although particular subgroups may occasionally show an effect. No further interpretation is warranted at this stage.

Correlations between E and N scales

In confirmation of the generalization made on the basis of previous research, all five correlations between E and N scales were negative in the present study (see Table 2), but the relationship was significant for Form B only, with $r = -0{\cdot}31$.

Correlations with age and intelligence

The correlations between EPI scales and age were obtained for the three large groups, and it was found that N scores decreased with age. The r values were $-0{\cdot}16$ (not sig.), $-0{\cdot}16$ ($P < 0{\cdot}05$) and $-0{\cdot}22$ ($P < 0{\cdot}01$) for groups 1, 2 and 3 respectively, using two-tailed tests. In all cases, both male and female subgroups gave negative correlations between the N scale and age. These results replicated those reported in the EPI manual. However, the negative correlations between the E scale and age reported in the manual were not replicated by the present studies, possibly owing to the more restricted age range of student groups: the standard deviations for age (see Table 1) were less than half the value given for the normal group in the manual.

When verbal intelligence was correlated with the other variables for group 1, the results differed for male and female subgroups. For females, no significant correlations were found; but for the 50 males, low IQ was associated with high Social Desirability responding and with low N (between IQ and D: $r = -0{\cdot}38$, $P < 0{\cdot}01$; IQ and L: $r = -0{\cdot}32$, $P < 0{\cdot}05$; IQ and N: $r = 0{\cdot}34$, $P < 0{\cdot}05$). It is possible, then, that the high D responding of the less intelligent males was what in part led to their having lower N scores.

DISCUSSION

It may be true that the EPI is 'a short, reliable and valid measure of the two most important dimensions of personality' (EPI manual). Nevertheless, the evidence reported here, together with previous research, demonstrates that the Inventory has certain characteristics which must be allowed for. In particular, the N scale is affected by Social Desirability response set, even when it is used with normal subjects for research purposes. The effect is pronounced when the scale is used for clinical assessment. At least some subjects, it would seem, wish to protect their self-esteem (Mosher, 1965), or to confound their assessors. Social desirability responding is not necessarily a result of deliberate faking or lying; indeed it has been suggested that normal subjects with high D scores are consistent in their self-descriptions, and careful to conform to test instructions as they interpret them (Heilbrun, 1965).

Blackburn (1965) considered that the MPI's 'value as a screening instrument is suspect', and this may well apply to the EPI also. Not only may some subjects obtain low N scores because of their high D responding, but in some circumstances *high* N scores might be obtained by a 'negative D' effect (Salas, 1968). The present authors agree with Huggan (1968) that 'considerable dubiety surrounds the use of scales like Eysenck's Neuroticism as a means of testing hypotheses derived from clinical experience'. Even when obtained in a 'pure' research situation, mean N values which are deviant from the normative value must be interpreted with care. For instance, the low N value of 7·76 obtained for male Australian medical students by Mowbray & Davies (1968) suggests that the students were tested under conditions giving rise to high D responding, rather than that they were highly stable.

A further implication concerns the validity of the N scale. Harrison & McLaughlin (1969) found a correlation of 0·56 between the EPI N scale and self-ratings of anxiety. But since both these variables may be similarly influenced by the D variable, the validity correlation is likely to be an overestimation of the 'true' value with the D variable controlled. Self-ratings, then, do not provide an appropriate criterion for validation of the N scale.

The technique for controlling D response set that is advanced in the EPI manual, of 'eliminating' high L scale subjects, is unsatisfactory. The proportions of subjects with L scores ≥ 4 were 16, 28, 9, 39 and 73 per cent for groups 1–5 respectively (cf. Koller & Castanos, 1969). Were such proportions of subjects to be rejected, research samples would be biased, and most renal donors would be lost merely because of their wish to seem worthy. In any case, such a procedure might reduce the effect of D responding on N scores, but would not eliminate it. (For example, the mean N score of the four donors with low L values of less than 4 was 7·5.) Indeed, there remained a significant correlation between N and L scales in group 1 even after the exclusion of all subjects with $L > 4$: $r = -0{\cdot}24$, with 84 subjects ($P < 0{\cdot}05$).

Another characteristic of EPI scales that must be allowed for is the lack of direct comparability between male and female scores, and between Form A and Form B scores. Furthermore, the negative correlation between E and N scales on Form B implies that the orthogonality of these variables cannot be assumed in factor-analytic studies.

The approximately zero correlations between the E scale (for both Forms A and B of the EPI) and the Crowne–Marlowe D scale do not support Farley's (1966) hypothesis that extraverts generally endorse fewer D responses because of their weaker social conditioning.

With regard to the negative correlation between IQ and the Edwards D scale with group 1 males, it may be noted that negative correlations with IQ were also found by

Feather (1967) for the Crowne–Marlowe D scale with students, and by McKerracher & Watson (1968) for the L scale with subnormal psychopaths.

We are indebted to Miss Julie Beare, who helped with the scoring of group 1, and to Mrs Maxine Shephard, who acted as a judge for discriminating E(imp.) and E(soc.) subscales.

REFERENCES

Adcock, C. J. (1965). A comparison of the concepts of Cattell and Eysenck. *Br. J. educ. Psychol.* **35**, 90–97.

Bailey, J. E. & Metcalfe, M. (1969). The MPI and the EPI: a comparative study on depressive patients. *Br. J. soc. clin. Psychol.* **8**, 50–54.

Blackburn, R. (1965). Denial-admission tendencies and the Maudsley Personality Inventory. *Br. J. soc. clin. Psychol.* **4**, 241–243.

Colston, M. A. (1969). The E and N scales of the EPI—some further Australian data. *Aust. Psychol.* **4**, 59–61.

Cramond, W. A. *et al.* (1967). Psychological screening of potential donors in a renal homotransplantation programme. *Br. J. Psychiat.* **113**, 1213–1221.

Cramond, W. A. *et al.* (1968). Psychological aspects of the management of chronic renal failure. *Br. med. J.* i, 539–543.

Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *J. consult. Psychol.* **24**, 349–354.

Edwards, A. L. (1957). *The Social Desirability Variable in Personality Assessment and Research.* New York: Dryden Press.

Eysenck, H. J. & Eysenck, S. B. G. (1964). *Manual of the Eysenck Personality Inventory.* London: University of London Press.

Eysenck, S. B. G. & Eysenck, H. J. (1963*a*). An experimental investigation of 'desirability' response set in a personality questionnaire. *Life Sci.* no. 5, 343–355.

Eysenck, S. B. G. & Eysenck, H. J. (1963*b*). On the dual nature of extraversion. *Br. J. soc. clin. Psychol.* **2**, 46–55.

Eysenck, S. B. G. & Eysenck, H. J. (1969). Scores on three personality variables as a function of age, sex and social class. *Br. J. soc. clin. Psychol.* **8**, 69–76.

Farley, F. H. (1966). Social desirability, extraversion, and neuroticism: a learning analysis. *J. Psychol.* **64**, 113–118.

Farley, F. H. (1967). On the independence of extraversion and neuroticism. *J. clin. Psychol.* **23**, 154–156.

Farley, F. H. (1968). Global self-ratings, the independence of questionnaire drive and anxiety, and social desirability variance. *Acta psychol.* **28**, 387–397.

Feather, N. T. (1967). Some personality correlates of external control. *Aust. J. Psychol.* **19**, 253–260.

Gibson, H. B. (1962). The Lie scale of the Maudsley Personality Inventory. *Acta psychol.* **20**, 18–23.

Harrison, N. W. & McLaughlin, R. J. (1969). Self-rating validation of the Eysenck Personality Inventory. *Br. J. soc. clin. Psychol.* **8**, 55–58.

Heilbrun, A. B. (1965). The social desirability variable: implications for test reliability and validity. *Educ. psychol. Measur.* **25**, 745–756.

Huggan, R. E. (1968). Neuroticism, distortion and objective manifestations of anxiety in males with malignant disease. *Br. J. soc. clin. Psychol.* **7**, 280–285.

Knowles, J. B. & Kreitman, N. (1965). The Eysenck Personality Inventory: some considerations. *Br. J. Psychiat.* **111**, 755–759.

Koller, K. M. & Castanos, J. (1969). The Eysenck Personality Inventory: Australian experience with normal and abnormal groups. *Aust. Psychol.* **4**, 48–53.

Lovell, V. R. (1967). The human use of personality tests: a dissenting view. *Am. Psychol.* **22**, 383–393.

McKerracher, D. W. & Watson, R. A. (1968). The Eysenck Personality Inventory in male and female subnormal psychopaths in a special security hospital. *Br. J. soc. clin. Psychol.* **7**, 295–302.

MOSHER, D. L. (1965). Approval motive and acceptance of 'fake' personality test interpretations which differ in favourability. *Psychol. Rep.* **17**, 395–402.

MOWBRAY, R. M. & DAVIES, B. M. (1968). Scores obtained by Australian medical students on the Eysenck Personality Inventory. *Aust. J. Psychol.* **20**, 59–61.

RICHARDSON, J. F. (1968). Correlations between the extraversion and neuroticism scales of the EPI. *Aust. J. Psychol.* **20**, 15–18.

SALAS, R. G. (1968). Fakeability of responses on the Eysenck Personality Inventory. *Aust. J. Psychol.* **20**, 55–57.

SALAS, R. G. & RICHARDSON, J. F. (1968). Some Australian data on forms A and B of the EPI. *Aust. J. Psychol.* **20**, 11–13.

SPARROW, N. H. & ROSS, J. (1964). The dual nature of extraversion: a replication. *Aust. J. Psychol.* **16**, 214–218.

WILLIAMS, C. (1969). Meaningfulness of the Eysenck Personality Inventory Lie scale. *Aust. Psychol.* **4**, 54–58.

Manuscript received 12 *August* 1969

Revised manuscript received 10 *March* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 55–60
With 3 text-figures
Printed in Great Britain

The Effect of Speed on Comprehension in Schizophrenia

By GABRIELLE MARCHBANKS AND MOYRA WILLIAMS

Fulbourn and Cambridge United Hospitals, Cambridge

A task involving the carrying out of spoken commands of varied length and frequency was devised and administered to schizophrenic subjects and non-psychotic patients under four conditions of speed. Schizophrenic subjects made more errors as the speed of the commands was increased. Patient controls made more errors when natural pauses were eliminated than under the fastest speed condition. Schizophrenic subjects repeated words from the commands as the speed of delivery increased; patient controls repeated less as the speed increased.

A relationship between the speed of stimulus input and performance by schizophrenic patients has frequently been reported (Salzinger *et al.*, 1966; Williams, 1966; Yates, 1966) and several theories have been put forward to account for it. Most of these are based on Broadbent's (1958) model of a filter preceding a limited channel processing system attached to a short-term memory store. The failure in schizophrenia can be accounted for by either a failure of the filter or a breakdown of the processing system, both of which result in overloading of the short-term memory store and failure to organize incoming stimuli in an economical way.

In another series of experiments and observations it has been shown that the spontaneous speech of normal subjects contains frequent pauses, and that these particularly precede words of high information (Goldman-Eisler, 1958*a*,*b*). A possibility has been raised by Williams (1965) that these pauses may be utilized also by the listener as signals to slacken the filter or clear the short-term memory store for important incoming information. The fact that pauses are often used consciously by actors as a means of stressing important utterances supports this possibility. Moreover, if such pauses are not made (as commonly happens when a speaker reads out a communication instead of delivering it spontaneously) the content of the discourse can become difficult to grasp.

Besides transitional pauses, Williams has also noted that in the English language a clustering of two or more consonants sometimes causes slight breaks in the stream of sound emission, and has suggested that these may also have a function for the listener. They may chunk or parse the stimuli into manageable units.

In this paper we report an experiment in which the speed of the given commands was controlled in two ways. One way was by increasing and decreasing the speed of words uttered in each command by mechanical means. The second was by editing a tape-recording of the commands in such a way as to eliminate all the speech pauses. The response by schizophrenic and non-psychotic patients was compared under these conditions.

METHOD

Material

A task was constructed consisting of 12 commands varying in length and complexity. The key words in each command were adjectives and nouns. The verbs, articles and conjunctions were nearly the same in every command and thus considered to be held constant.

In six of the commands all the adjectives and nouns were taken from the AA words in the Thorndike & Lorge (1944) word count. In the other six commands the nouns used had frequencies from 1 to 10, and the adjectives had frequencies from 1 to 49.

In each group of six frequent and six rare commands, one command had one noun, one had one adjective and one noun, one had two nouns, one had two adjectives and two nouns, one had three nouns, and one had three adjectives and three nouns.

The commands were as follows:

Frequent commands: (*a*) Point to the dog. (*b*) Point to the big shoe. (*c*) Point to the chair and bed. (*d*) Put the two cats next to the small cup. (*e*) Put the chair, the iron and the bed together. (*f*) Put the green car, the two cats and the big shoe together.

Rare commands: (*g*) Point to the ant-eater. (*h*) Point to the pink thimble. (*i*) Put the cooker next to the motorcycle. (*j*) Point to the purple shoestring and the white thimble. (*k*) Put the toilet, bathtub and wardrobe together. (*l*) Put the yellow shoestring, the large scissors and the blue thimble together.

These commands were put into the following random order for presentation: *b*, *h*, *a*, *i*, *c*, *f*, *k*, *j*, *g*, *e*, *l*, *d*.

Miniature objects were spread out on a table in front of the subject. The objects presented were such that for correct performance each adjective and each noun had to be comprehended in each command. The objects were: 1 big shoe, 1 small shoe, 1 pink thimble, 1 white thimble, 1 blue thimble, 1 dog, 1 iron, 1 cooker, 1 motorcycle, 1 chair, 1 bed, 1 man, 3 cats, 1 large scissors, 1 small scissors, 1 toilet, 1 bathtub, 1 wardrobe, 1 purple shoestring, 1 yellow shoestring, 1 ant-eater, 1 small cup, 1 large cup.

The commands were read into a tape in a standard English female voice. The tape was put through a Mark 11 Information Rate Changer by courtesy of the Gotham Audio Development Corp., New York. This is a device which can speed up or slow down a tape-recording without distorting the pitch. By this process two new tapes were made: one which speeded and one which slowed each command by 35 per cent.

Another copy of the original tape was edited by courtesy of the BBC Editing Department. By putting the tape very slowly through the head of the recorder containing built-in splicing equipment, the editor was able to remove the natural pauses in each command. The original and edited tapes were played through an oscilloscope, and it was possible to see that in the edited tape, the areas of silence had been virtually eliminated.

Subjects and procedure

Forty chronic and 40 acute patients diagnosed by the psychiatric clinicians as suffering from schizophrenia were divided into four groups (each group receiving the task under a different condition). The acute patients were in their first or second admission to a mental hospital. The chronic patients had either been living in the hospital for at least 3 years or had been admitted at least four times previously with diagnoses of schizophrenia. The 80 subjects were subdivided into 'thought-disordered' or 'non-thought-disordered', based on an interview during which the subject was asked to explain the meaning of four proverbs and to answer an open-ended question. The 80 subjects were divided into four groups of 20, each group being given the task under a different condition as follows.

Group 1 (*ordinary speed*): (*a*) Ten acute schizophrenic patients (five men, five women; mean age 33·7 years, ranging from 16 to 57 years); five were thought-disordered, four were non-thought-disordered, one was undetermined. (*b*) Ten chronic schizophrenic patients (five men, five women; mean age 36·3 years, ranging from 19 to 59 years); four were thought-disordered, four were non-thought-disordered, two were undetermined.

Group 2 (*slow speed*): (*a*) Ten acute schizophrenic patients (five men, five women; mean age 35·1 years, ranging from 17 to 55 years); five were thought-disordered, four were non-thought-disordered, one was undetermined. (*b*) Ten chronic schizophrenic patients (five men, five women; mean age 36·9 years, ranging from 21 to 60 years); five were thought-disordered, three were non-thought-disordered, two were undetermined.

Group 3 (*fast speed*): (*a*) Ten acute schizophrenic patients (five men, five women; mean age 34·5 years, ranging from 16 to 50 years); two were thought-disordered, four were non-thought-disordered, four were undetermined. (*b*) Ten chronic schizophrenic patients (five men, five women; mean age 38·4 years, ranging from 21 to 60 years); three were thought-disordered, two were non-thought-disordered, five were undetermined.

Group 4 (*edited speed*): (*a*) Ten acute schizophrenic patients (five men, five women; mean age 37·2 years, ranging from 17 to 59 years); seven were thought-disordered, two were non-thought-disordered, one was undetermined. (*b*) Ten chronic schizophrenic patients (five men, five women; mean age 35·9 years, ranging from 16 to 55 years); six were thought-disordered, three were non-thought-disordered, one was undetermined.

The control subjects were 40 non-psychotic psychiatric patients who were randomly assigned to one of the four groups, except that an equal number of men and women were in each group. Their diagnoses were: reactive depression (10 subjects), inadequate personality (10 subjects), phobic state (two subjects), obsessional (four subjects), identity crises (two subjects), personality disorder (eight subjects) and anxiety (four subjects).

The subjects were seated in front of the miniature objects and were told: 'Here are some toys. I will play a tape-recording on which there are instructions telling you what to do with the toys. For example, you might hear the tape say "point to the man" and then you would point to the man. [The experimenter points to the man, by way of demonstration.] Now I will give you an example instruction and you do it. Put the man next to the chair. [The subject does this.] Now I will put on the tape, and you will follow the instructions.' Before each of the conditions in which the speed was distorted, the experimenter added: 'The speed of the speech on the tape has been changed so that it is faster (slower) than usual.'

The experimenter recorded (1) whether or not the command was carried out correctly, and (2) the number of times subjects repeated key words (i.e. adjectives and/or nouns) out loud as they conducted their search for the response.

RESULTS

Effect of speed on errors. In Fig. 1 the mean number of errors for schizophrenic and non-psychotic control patients shows that the fast condition was the most difficult for schizophrenic patients, but the edited condition was the most difficult for the patient control group.

For schizophrenic subjects the slow and ordinary conditions were not significantly different, but the ordinary and fast conditions were different ($P < 0{\cdot}02$; Mann–Whitney test, two-tailed). The fast and edited conditions were not significantly different, but neither were the ordinary and edited conditions.

Fig. 2 demonstrates that the fast condition was the most difficult for each schizophrenic subgroup.

In the control group, the fast and ordinary conditions and the fast and edited conditions were significantly different ($P < 0{\cdot}05$; Mann–Whitney U test, two-tailed).

Effect of speed on amount of repeating. The mean number of repeats and errors go hand in hand for schizophrenic subjects in general so that the more errors there were, the more repeats had been made (see Fig. 1). The opposite was true for control subjects in the slow, ordinary and fast conditions. The more errors controls made, the less they repeated, except in the edited condition.

When the schizophrenic subgroups are considered separately, some curious discrepancies appear. If the division is made into chronic and acute, both groups make most repeats in the most difficult condition, i.e. the fast condition (Fig. 2).

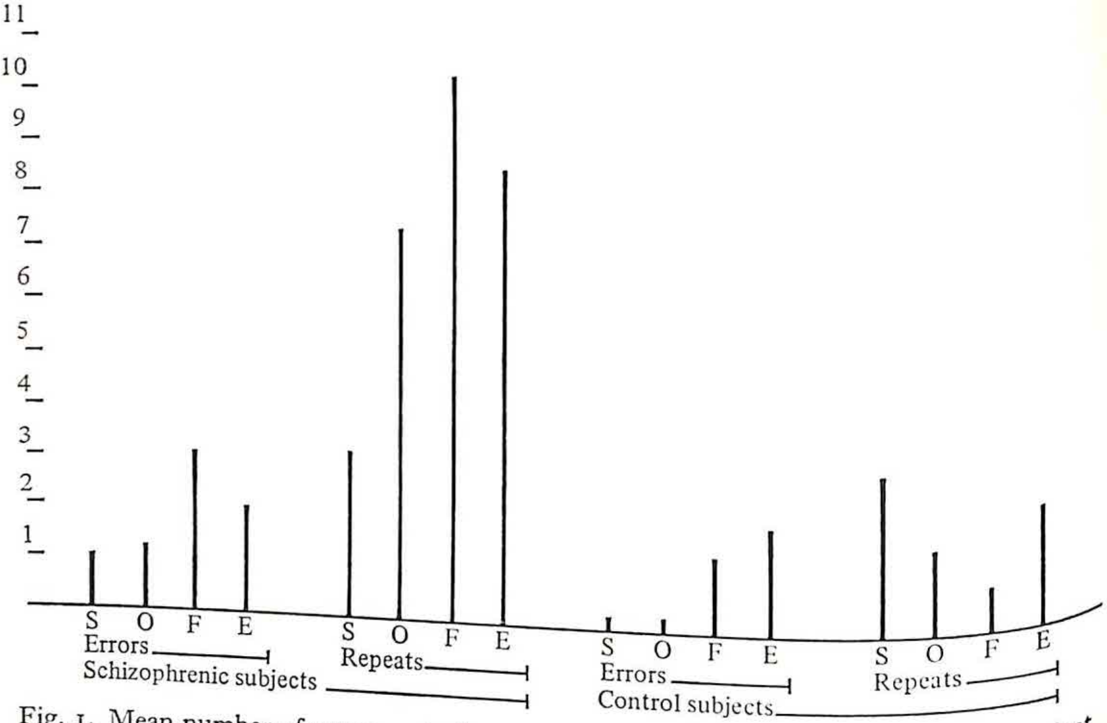


Fig. 1. Mean number of errors and of repeats made by schizophrenic subjects and patient control subjects in four speed conditions. S, slow condition; O, ordinary condition; F, fast condition; E, edited condition.

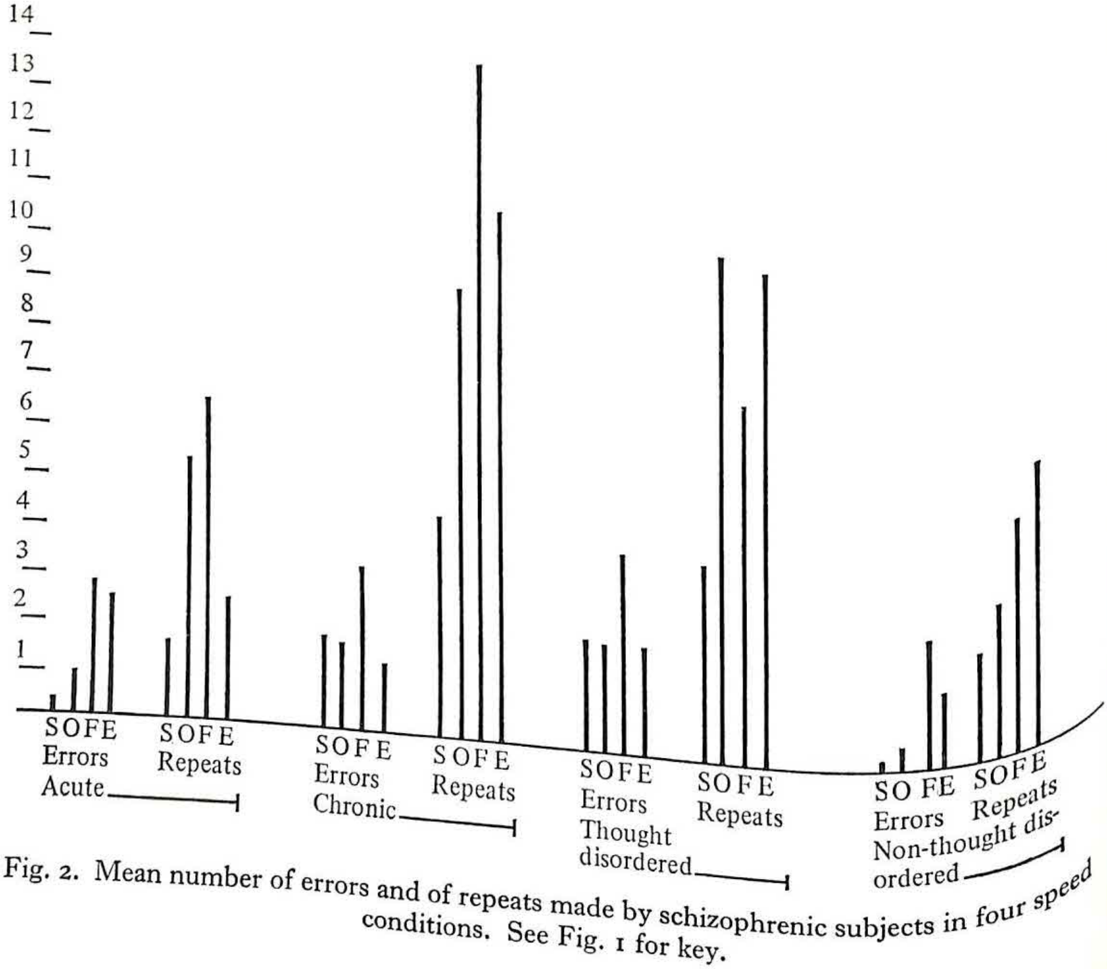


Fig. 2. Mean number of errors and of repeats made by schizophrenic subjects in four speed conditions. See Fig. 1 for key.

When these subjects were subdivided into thought-disordered and non-thought-disordered groups, however, this did not hold true (see Fig. 2). The non-thought-disordered subjects made most repeats in the edited condition and the thought-disordered group made most repeats under the ordinary and edited conditions.

Effect of length of command on errors. In the slow, ordinary and fast conditions there appears to be a general trend for more errors to be made to longer commands than to shorter commands by both schizophrenics and controls. In the edited condition the longer commands were more difficult for the control group, but the schizophrenic group made more errors in the middle length commands (see Fig. 3).

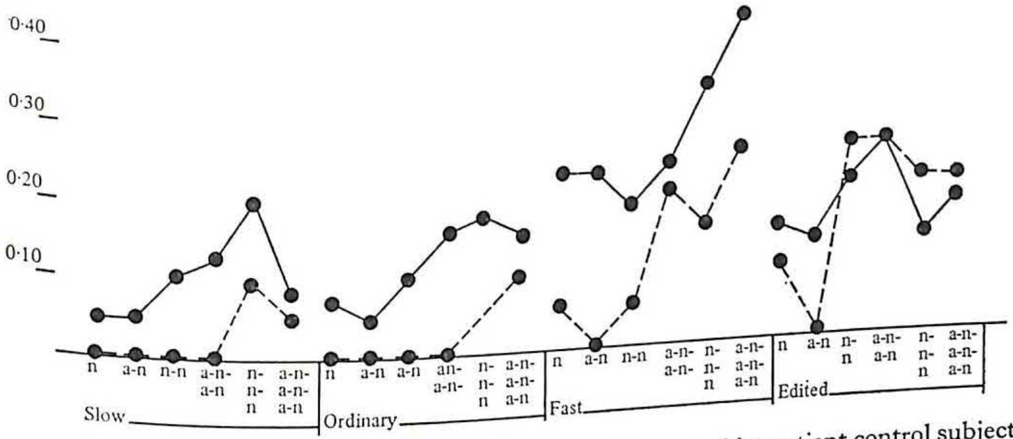


Fig. 3. Proportions of errors made by schizophrenic subjects and by patient control subjects to commands of different lengths in four speed conditions. ——, schizophrenic subjects; ----, patient control subjects; n, one noun; n–n, two nouns; n–n–n, three nouns; a–n, adjective and noun; a–n–a–n, two adjectives and nouns; a–n–a–n–a–n, three adjectives and nouns.

It appears that the addition of adjectives to nouns did not inevitably increase errors. Indeed, in some cases, it seemed to make the command easier to obey. Adding nouns, however, did increase errors by schizophrenics under the slow and ordinary conditions.

Frequent and rare commands. No differences between errors made to the frequent or rare commands were found under any of the conditions for schizophrenic or control subjects.

DISCUSSION

In the above experiment it was found that increasing the speed at which commands were given increased the errors of both control and schizophrenic patients, although schizophrenics made more errors under the slow, ordinary and fast conditions than control subjects. However, under the edited condition it was found that the number of errors was very similar in the two groups.

Since it has been frequently found that schizophrenic patients perform worse than patient control groups (McGhie, 1969) this might be interpreted as indicating that the edited condition was relatively more difficult for control subjects than for schizophrenic subjects. It might be concluded from this that schizophrenics do not utilize natural speech pauses in the same way as controls, and this may be one factor hampering their ability to comprehend.

Possibly some clues as to the function of these pauses may be obtained from a study of the repetitions. The control group, matching their rate of responding to the rate of the stimulus presentation, repeated items less as the speed of the input increased except when the speech pauses were eliminated; then their number of repetitions increased. The schizophrenic subjects, on the other hand, appeared to be trying to delay the rate of stimulus presentation by increasing their repetitions as the stimulus rate increased. It appears from this that repeating may be a strategy used to retain stimuli in the short-term memory store until they have been adequately processed, and that while normal subjects carry out this process so fast that at the slow speeds a word may be forgotten before the next one arrives unless it is repeated, schizophrenics process much more slowly.

The fact that the performance of the schizophrenic patients did not deteriorate even further when the natural speech pauses were eliminated in the edited condition, suggests that one possible reason for the schizophrenics' slowness in processing may be that at each of these pauses their short-term memory store becomes overloaded with material from some other (matching) source, so blocking the arrival of new material.

It has already been shown in a previous communication that schizophrenic patients have difficulty in word identification when associated material is presented to them; conceivably under the edited condition in this experiment there was no time for the associated material from their word store to be called up into the short-term memory.

We are indebted to the Mental Health Research Fund for providing a grant to make this research possible. We also wish to thank the consultants of Fulbourn and the United Cambridge Hospitals for permission to study their patients.

REFERENCES

BROADBENT, D. (1958). *Perception and Communication*. Oxford: Pergamon Press.
GOLDMAN-EISLER, F. (1958*a*). The predictability of words in context and the length of pauses in speech. *Lang. & Speech* **1**, 226–231.
GOLDMAN-EISLER, F. (1958*b*). Speech production and the predictability of words in context. *Q.J. exp. Psychol.* **10**, 96–106.
MCGHIE, A. (1969). *Pathology of Attention*. Harmondsworth: Penguin Books.
SALZINGER, K., PORTNOY, S. & FELDMAN, R. (1966). Verbal behaviour of schizophrenic and normal subjects. *Ann. N.Y. Acad. Sci.* **105**, 845–860.
THORNDIKE, E. & LORGE, I. (1944). *Teacher's Word Book of 30,000 Words*. New York: Columbia University Press.
WILLIAMS, M. (1965). The effect of pauses on the interaction of speech. (PLU/65/24 MRC Psycholinguistics Research Unit. Unpublished.)
WILLIAMS, M. (1966). The effect of context on schizophrenic speech. *Br.J. soc. clin. Psychol.* **5**, 161–171.
YATES, A. (1966). Psychological deficit. *Ann. Rev. Psychol.* **17**, 111–144.

Br. J. soc. clin. Psychol. (1971), **10**, 61–72
Printed in Great Britain

High Blood Pressure as a Psychosomatic Disorder: A Selective Review

By RAYMOND COCHRANE

Michigan State University, East Lansing, Michigan, U.S.A.

A selective review of medical, psychological and sociological studies of high blood pressure is made, the criterion for inclusion being that hypertension is approached as a psychosomatic disorder. The enormous amount of literature produced to date yields general support to the idea of a link between perceived stress in the environment, a personality overreactive to stress and high blood pressure. The specific mechanisms involved have not as yet been clearly identified. The literature is reviewed in the areas of personality and hypertension, experimental elevation of blood pressure, environmental factors in hypertension and explanations of high blood pressure.

This paper is confined to reviewing those studies which have looked at the evidence for hypertension (or at least some forms of it) being of psychosomatic origin. For the purposes of this paper 'psychosomatic' is used to describe the general approach to medical and psychiatric problems which recognizes and appreciates that ostensibly physical ailments exist within a psychological context. The use of the term 'psychosomatic' implies that psychological and social factors can operate in a causative manner in the production of physical illness.

Four criteria are established which have to be met before hypertension can be considered of psychosomatic origin. First, what is known about the physiology of hypertension must not be incompatible with the psychosomatic explanation. Secondly, it would strengthen the case if hypertension could be produced experimentally by manipulating psychological variables. Thirdly, a relationship between environmental factors and the onset of hypertension would strengthen the psychosomatic explanation. Last, there should be some measurable difference between individuals who are susceptible to the illness and those who are not—in this case personality differences are expected. Literature is reviewed in each of these areas.

THE PHYSIOLOGICAL BASIS OF HYPERTENSION

It is not the intention here to attempt a thorough review of medical thinking on mechanisms which may produce essential hypertension. From the extensive reviews of the medical literature published by Goldberger (1958), Simonson & Brozek (1959) and Geiger & Scotch (1963) two things emerge. First, that the principal cause of essential hypertension remains unknown. Secondly, what is known about the physiological correlates of hypertension is consistent with regarding it as a psychosomatic disorder. Certainly the autonomic nervous system is implicated in hypertension and in a wide variety of symptoms which are characteristic of neurosis (Kuntz, 1953). Mills (1959) and Shapiro (1960) among others have suggested that a chain of events beginning with emotional stress and ending with hypertension is quite probable. The intermediate links seem to be increased noradrenalin secretion produced by sympathetic nervous system activity, and consequent renal vasoconstriction and increased renin output. Apparently increased renin output from the kidneys may be converted to Angiotensin II, which eventually stimulates the adrenal cortex to secrete aldosterone, and this leads to marked sodium retention (Mendlowitz *et al.*, 1962; Genest *et al.*, 1961). The significance of this is that sodium retention in the arteriolar muscle cells sensitizes the vascular smooth muscle. The hyper-reactive blood vessels respond to normal neural firing with vasoconstriction and functional hypertension.

While it cannot be maintained that increased renin output from the kidneys is the sole source of hypertension (Ledingham, 1951, and others have produced post-nephrectomy hypertension in rats) the weight of evidence is sufficient to provide a convincing case that stress and hypertension are related through the autonomic nervous system.

EXPERIMENTAL ELEVATION OF BLOOD PRESSURE

Stress and autonomic lability

Autonomic reaction to stress is well known. Although there is a considerable amount of conflicting evidence, several studies have reported organic deterioration in animals caused by prolonged autonomic stimulation (Brady, 1958; Sawrey & Weisz, 1956; Sines, 1962; Hudak & Buckley, 1961).

Given that chronic stressful stimulation produces autonomic responses and eventually organic deterioration, it remains to be explained why the specific reaction pattern displayed differs from individual to individual—some developing ulcers, others asthma or rheumatism and others hypertension.

The work of Lacey (1950, 1953, 1955, 1956, 1958) indicates that individuals exhibit idiosyncratic patterns of somatic reaction to mild stress. This finding has important consequences for the psychosomatic interpretation of certain illnesses when reviewed in conjunction with two other findings. Lacey & Smith (1954) showed that rapid unconscious generalization of conditioned anxiety responses can occur. Their subjects were required to chain-associate to a set series of stimulus words for 15 sec. At the end of one of the chain associations (i.e. those following the stimulus word 'cow') they were shocked. Subjects very rapidly developed autonomic changes upon hearing and responding to the word 'cow'. As far as could be determined this was an unconsciously formed association; even after the experiment the subjects were not aware that the 'cow' chain association had uniquely and consistently been associated with shock. In the same experiment it was noticed that not only did the word 'cow' acquire the property of being able to induce autonomic activity (in this case heart rate) but generalizes to other related, agricultural words (plough, corn, harvest, chicken, tractor, etc.) and unlike other studies of generalization they found that the generalized response was of greater magnitude than the primary response. Also there appeared to be a relationship between the existing level of 'chronic anxiety' in the individual (as measured by the

Manifest Anxiety Scale) and the ease with which conditioned anxiety was acquired and generalized.

The second line of investigation, led by Malmo (1957), discovered that when individuals with various psychosomatic complaints are subjected to stress they are most likely to respond maximally with disturbances in the 'critical symptom areas', i.e. the area in which the psychosomatic complaint was located. This gave rise to the concept of 'symptom stereotype' which was described by Lacey & Lacey (1958) as the 'constancy of the physiological measure in which maximal activation being consonant with the somatic complaint'.

A striking example of autonomic response specificity was reported by Engel & Bickford (1961). They used 20 female hypertensives and 20 matched controls and subjected them to five stimuli while taking measurements of physiological response functions. They found no difference in the frequency of occurrence of individual response specificity between the two groups but 15 of the 20 hypertensives showed their greatest degree of response in blood pressure, while only five of the 20 normotensives did. They found no differences in any indices of generalized arousal or total activation (i.e. the sum of all autonomic responding), the only important difference being in the patterning of the responses. They interpret their results to mean that hypertensives are generally over-reactive to environmental stressors in blood pressure.

None of the authors quoted felt it necessary to speculate on how individual response patterning developed but it could either be due to genetically determined factors or to the chance acquisition of patterns by individuals according to the ordinary laws of learning. It is possible that psychosomatic disorders are the result of prolonged over-reactivity in one specific area which is enhanced by massive generalization of stress to numerous previously neutral stimuli.

Experimentally produced hypertension

Many of the studies referred to in the previous section have incidentally shown that blood pressure consistently rises when an individual is stressed—this being more marked in the case of hypertensives where it produces a greater relative rise in blood pressure than in normotensives. McGinn *et al.* (1964) have reviewed this literature in considerable detail.

There have been many attempts at producing hypertension in previously normal animals by subjecting them to stress. For obvious reasons this has not been attempted with humans. In 1945 Medoff & Bongiovanni produced hypertension in rats by subjecting them to air blasting for 10 min. a day. By the time the animals were 400 days old, 61 per cent of the experimental animals were hypertensive and only 19 per cent of the control group. They also found that the experimental animals that developed hypertension showed higher behavioural reaction to stress than experimental animals that did not develop hypertension. There were, however, several weaknesses in this research; the rats were anaesthetized while blood pressure was measured, and some of the animals died during the course of the experiment. Hudak & Buckley (1961), using extremely stressful procedures (flashing spot lights, 100 db auditory stimulation and rapid oscillation), produced hypertension in rats, but 15 of the 24 experimental animals died before the 35-week experimental period was over. Simonson & Brozek (1959) reported on Russian work which has produced stable hypertension in rats by continued painful stimulation. Shapiro & Melhado (1958) failed to produce hypertension or any other organic pathology in rats subjected to prolonged stress, but there was evidence that existing hypertensive vascular disease could be aggravated.

A particularly relevant experiment is that of Farris *et al.* (1945). They obtained a measure of emotionality (from Hall's open-field situation) and blood pressure before subjecting the experimental animals to stress (the sound of air blasting) and then repeated the blood pressure measurements. They found that all the initial blood pressure measurements were normal and that the highly emotional animals who were subjected to stress became hypertensives whereas those with low emotionality scores did not. In the control group, which contained animals matched for emotionality scores, hypertension did not develop. Apart from some deficiencies in reporting this experiment, it can be criticized because no attempt was made to discover whether the post-stress measurements represented a temporary elevation of blood pressure possibly as a direct response to the stress, or stable hypertension. Maher (1966) reports work done by Miminoshvili in Russia in which a deliberately established approach–avoidance conflict in monkeys eventually produced hypertension. In fact, Simonson & Brozek quote many studies in which the establishment of experimental neuroses was accompanied by the onset of hypertension, although this was not usually stable (i.e. it ended when the conflict producing situation was changed).

Russian workers have also led the field in the production of conditioned high blood pressure. Napalkov & Karas (reported by Simonson & Brozek) found that strong unconditioned stimuli (e.g. pistol shots) produced a significant transient increase in blood pressure in dogs. When conditioned signals were associated with these they produced a greater increase of blood pressure than the unconditioned stimuli and this exaggerated response was not susceptible to extinction in the normal way. A stable hypertension developed which persisted until observation ceased. In the same paper Simonson & Brozek report the successful use of adrenalin injection as the unconditioned stimulus and sound as the conditioned stimulus producing a rise in blood pressure and other features associated with essential hypertension. These dogs were in fact producing a rise in blood pressure as a conditioned response with no conflict or painful stress but the pharmacological elevation of blood pressure. If this were demonstrated in humans it would mean that previously neutral stimuli which had 'accidentally' been temporally related to an increase in adrenalin secretion would acquire the property of being able to raise blood pressure.

Dykman & Gantt (1960) succeeded in producing specific blood pressure rises to three tones which had previously been associated with shock in dogs. This conditioned hypertension was quicker to form and more persistent than motor conditioned reflexes, thus confirming work reported previously. The most significant finding from this experiment (unfortunately $n=4$) is that even after more than a year's rest during which the dogs had received no further conditioning the specific responses were retained and in one case the conditioned hypertension had become exaggerated, showing that there had been some organic development.

The conclusion is inescapable that hypertension can be produced in animals as a direct result of psychological stress.

ENVIRONMENTAL FACTORS IN HYPERTENSION

It has frequently been reported that environmental stress raises blood pressure temporarily, both in animals (Wilhelmj *et al.*, 1953) and in man (Van der Valk, 1957; Stevenson *et al.*, 1952). Temporary blood pressure elevation cannot be equated with hypertension, however. If the life-stress hypothesis is to be tested, then it is more productive to approach the problem retrospectively.

Natural experiments can be taken advantage of. Two widely quoted examples are the work of Graham (1945), who found that a quarter of soldiers with long combat experience were hypertensive for several months afterwards, and Ruskin *et al.* (1948), who found that over one half of the survivors of an earthquake in Texas City were hypertensive two weeks after the event.

In a prolonged series of studies in New York, Harold Wolff and his colleagues have attempted to explore in detail the interaction between life situation, emotion and blood pressure (Wolf *et al.*, 1948, 1955). Of particular interest were some 'ecological' studies involving groups or small populations. 'Each of these groups was selected because its members have shared some similar experience in their environment, and each individual in each group has therefore participated in an experiment of nature which has been replicated many times' (Hinkle, 1961). They studied six groups for periods ranging from 11 to 25 years. Hinkle stresses repeatedly that men react to their life situations as they perceive them, so that individuals sharing identical life situations may nevertheless experience them in entirely different ways. They found that among Hungarian refugees and a Chinese sample hypertension was much more common than in a group of American working people: 'for the American working people, by and large, had spent their lives in a relatively unchanging and secure environment, whereas the Chinese and Hungarians, being refugees and migrants, had been exposed to profound changes in their social and physical circumstances, and they were now engaged in trying to make new lives for themselves in a society with which they were quite unfamiliar.' Hinkle & Wolff (1961) followed up this suggestion with a survey of American 'first level managers'. They divided this group into college graduates who had had a 'lifetime of familiarity' with the social environment in which they worked, and managers who had risen from the ranks who were 'living and working in a social environment different from their early environment, which they perceived as new, unfamiliar, and full of challenges'. Not only did the latter group have more illnesses of all kinds during the one year observation period they showed more signs often considered prognostic of cardiovascular disease. They again concluded that these differences were due to the way in which these two groups perceived objectively quite similar present life situations. 'Hypertension thus appears as one of the ways that threatened, frustrated and overburdened men can react to their experiences.'

However, therapeutic efforts based on this hypothesis had no major success in reversing the course of the disease in hypertensives.

More substantial evidence is provided by Cruz-Coke (1960). Studying Peruvian policemen, he showed the blood pressure of 21 men who left their native village and emigrated to westernized Lima was higher than that of a comparable sample who remained.

Morris *et al.* (1953), in an epidemiological study of coronary heart disease, found there was a 'cluster of unfavourable factors among certain occupations, particularly in the upper social classes'. Their main hypothesis was that men in physically active jobs have less heart disease in middle age than have men in physically inactive jobs. In a later study (Morrison & Morris, 1959) it was found that there was a higher incidence of hypertension among bus drivers than among bus conductors. Bus drivers are both less active and more subjected to stress than bus conductors.

In the United States no consistent differences between occupational groups in hypertension have been found (Scotch & Geiger, 1963). In the Soviet Union work involving 'nervous tension' has been shown several times to produce more hypertensives than jobs involving physical work (Simonson & Brozek, 1959).

At present one is forced to agree with Glock & Lennard (1957): 'It is evident that in its present form the stress hypothesis does not contribute in any precise way to our understanding of the causes of hypertension.'

PERSONALITY AND HYPERTENSION

Of the enormous variety of personality traits which have been postulated as being associated with hypertension, few have received any consistent empirical support (Glock & Lennard, 1957). Alexander (1939) originated the 'repressed hostility' hypothesis. In essence Alexander's hypothesis is that one neurotic way of handling a conflict between aggressive impulses and social control is to repress these impulses, whereupon they provide a constant source of irritation which produces a chronic elevation of blood pressure. Hypertensives should therefore differ from normals in two ways. They should exhibit a higher level of repressed hostility. They should also be more neurotic because it is only those who fail to manage their hostile impulses successfully who repress them.

Many of the psychoanalytically orientated workers in this area have confirmed Alexander's hypothesis. Binger (1945) is a typical example. Unfortunately the quality of the measurement involved in these studies has not always been sufficiently high for the findings to be acceptable to experimental psychologists. The clinical interview unsupported by any other evidence is generally considered to be of dubious validity.

Harburg *et al.* (1965), using somewhat more objective measures, were unable to confirm the hypothesis of the hypertensive personality. From a review of the literature they found that the hypertensive was typically described as anxious, sensitive, inwardly hostile, but submissive in social situations, this resulting, in part, from a threatening parent–child relationship. Using a questionnaire and the Semantic Differential on healthy male college students they found no support for the hypothesis that subjects with higher blood pressure levels recall having punitive, irritable, unloving parents. However, blood pressure variability was associated with a 'stern dominant father, perhaps overly concerned with mobility (social)'.

Other relevant studies on the repressed hostility hypothesis include Harris (1953), who studied U.S.A.F. officers by prolonged observation and interview and found that those with high blood pressure were generally similar to others but were said to be more rigid, over-controlling of impulse, less resilient and less flexible and more likely to have been exposed to pathogenic events in childhood. Kalis *et al.* (1961) make the important point that if psychological factors influence the development of hypertension and are not merely incidental or secondary, then they should exist before the onset of the disease. They used a group of young adults with the 'predisposition to hypertension' (i.e. those with a temporarily raised blood pressure). The results of an analysis of adjective checklists, rating scales and Q-sort descriptions following life history interviews showed that pre-hypertensives are more anxious, have more conflict in the area of identity and 'difficulty in the control and expression of hostility' than a control group. They interpret their findings in the framework of psychoanalytic theory and conclude that the prehypertensive is ill-equipped to meet life stress because of ego defects resulting from developmental disturbances.

An otherwise excellent study by Gressel *et al.* (1949) suffered from the common limitation that the only personality data available were from a series of clinical interviews. They used 50 hypertensives and two matched control groups who had either one other psychosomatic complaint or a chronic medical–surgical disorder. All personality ratings used (hysteria, anxiety, depression, obsessive-compulsive behaviour, subnormal assertiveness and impulsiveness), except impulsiveness, differentiated hypertensives from those with chronic organic disorders. The two control groups were also easily distinguishable on the rating scales. When the hypertensives and those with personality disorders were compared only the scores on obsessive-compulsive behaviour and subnormal assertiveness were different. A further interesting finding is that no personality differences were found between hypertensives classified as 'neurogenic' and others.

Stokvis (1953) also compared hypertensives with other psychosomatic groups (including asthma and ulcerative colitis). He suggests that although in general all psychosomatic patients display infantile personality structure which is not specific to their affection, those in the hypertension–migraine–cardiac neurosis group are more likely to exhibit repression of aggressive impulses.

An important series of studies has been reported by Hokanson and colleagues (1961, 1962*a*, *b*, 1963). Normal subjects were deliberately made angry by the experimenter while their blood pressure was being recorded. An elevated pressure could be brought back to normal by giving the subjects the opportunity to administer electric shock to the experimenter (i.e. some overt expression of hostility). The blood pressure of those subjects who were angered but not allowed to express their feelings remained higher. If the anger-provoking person was of obviously higher status, blood pressure quickly returned to normal even without an opportunity to act aggressively, suggesting that the frustration of the expectation that aggression may be allowed is the key factor. Holmes (1966) reports findings that appear opposite to the major findings reported in the Hokanson studies. Frustrated subjects who were given an opportunity to aggress reduced their level of physiological arousal (including blood pressure) *less* than those who were not permitted to aggress. Holmes suggests that his experimental situation differs from Hokanson's in that the provocateur left the situation and therefore there were fewer cues available to make it easy for the subjects' aggression to perseverate. Results reported by Gambaro & Rabin (1969) emphasize that other personality factors interact with the opportunity to express aggression. Generally they found that 'low-guilt' subjects (on the Mosher Incomplete Sentences Test) exhibited a significantly greater decrease in blood pressure when given the opportunity to aggress after being angered than did 'high-guilt' subjects. Attempts such as these to operationalize clinical insights are potentially the source of the most valuable evidence on the repressed aggression hypothesis.

A recent study in India (Ghosh & Bashey, 1966), using matched control groups, employed a wide variety of proven psychological tests together with an unstructured interview. Twenty-five patients with essential hypertension were matched by age and sex with 25 carcinoma patients. The tests used included the Rorschach, Taylor's Manifest Anxiety Scale and part of the Cornell Medical Index. The carcinoma patients had a significantly higher anxiety level than the hypertensives but were much better at expressing their anger and hostility. Ghosh & Bashey say that the anxiety of the carcinoma patients seemed to be closely related to their present diseased condition and was not necessarily indicative of their basic personality structure. Nonetheless the psychological tests used indicate that the hypertensives were less anxious on the MAS. This particular study lends support to one of Alexander's hypotheses, namely the repression of hostility, but contradicts the idea that the hypertensive patient suffers from emotional lability.

The results of studies employing measures of neuroticism–anxiety–emotional lability are contradictory and it is perhaps no coincidence that in this type of investigation objective tests have been more widely used. The studies of repressed hostility have frequently used subjective assessment and have yielded consistent results.

Harburg *et al.* (1964) obtained scores on the Cattell 16 PF test and the Mandler-Sarason Test Anxiety Questionnaire from 74 male college students together with several blood pressure measurements. On measures of anxiety they found that students who had high initial blood pressure readings which eventually subsided scored higher. The group that showed evidence of sustained high blood pressure could not be statistically distinguished from normotensives. Product moment correlation coefficients of +0·21 and +0·26 between systolic blood pressure and scores on the Neurotic and Anxiety scales of the 16 PF are reported. It is possible therefore that anxiety and neuroticism are more closely associated with a tendency to greater blood pressure lability (which could explain a high initial reading at college registration) than with sustained hypertension.

Using the MPI in Britain, Sainsbury (1960) found that hypertensives, in common with other psychosomatic out-patients at two general hospitals, scored higher on the Neuroticism scale than a control group of medical out-patients. It was confirmed by Robinson (1963) that out-patient hypertensives have higher N scores than a sample with high blood pressure from outside hospital, but of great interest is his finding that within the population outside hospital there was no relationship between blood pressure and neuroticism. Not only was the overall correlation between blood pressure and neuroticism insignificant but high and low blood pressure groups chosen from this sample could not be distinguished on neuroticism.

Cochrane (1969) found no difference between the neuroticism scores of patients discovered by their general practitioner to be hypertensive and a matched control group being treated for other reasons. Both of these groups had higher neuroticism scores than a group not currently being treated. It is possible that some of the drug treatments being used at the time Sainsbury and Robinson did their studies may have artificially elevated neuroticism scores.

In summary the two most significant hypotheses about the personality of hypertensives are that: (*a*) hypertension develops in individuals who are emotionally labile and have neurotic tendencies, and (*b*) repression of hostile and aggressive impulses is characteristic of the hypertensive personality.

While the findings on the former hypothesis are inconclusive it appears to be worth following up; the second hypothesis has been confirmed repeatedly but in many cases the quality of the evidence leaves something to be desired.

CONCLUSIONS

The specific conclusions drawn from this review are:

1. Physiological explanations of blood pressure elevation are consistent with the possibility that some forms of essential hypertension have psychosomatic origins.
2. Elevation of blood pressure can be conditioned to previously neutral stimuli and this is resistant to extinction, and may in fact become exaggerated after conditioning has ceased.
3. Long-lasting changes in blood pressure can be produced in man and many different species of animal by noxious stimulation.
4. There is some evidence that hypertension develops in individuals who are emotionally labile, have neurotic tendencies and repress hostile and aggressive impulses.
5. Despite the enormous amount of energy expended in attempting to find consistent correlates of high blood pressure over the last three decades very few factors have been identified as being conclusively implicated in hypertension. It seems probable that a complex interaction between felt stress in the environment, personality and personal history contributes to the development of hypertension in some individuals. This, however, is unlikely to completely account for the phenomenon of essential hypertension.

This work was conducted while the author was supported by Medical Research Council Grant G/964/401/C awarded to Dr J. O. Robinson, University College, Cardiff. The author

wishes to thank Dr Robinson for his help in planning the review and reading the manuscript.

REFERENCES

ALEXANDER, F. (1939). Emotional factors in essential hypertension: presentation of a tentative hypothesis. *Psychosom. Med.* **1**, 173–179.

BINGER, C. A. *et al.* (1945). Personality in arterial hypertension. Am. Soc. for Res. in Psychosom. Monogr. no. 8.

BRADY, J. V. (1958). Ulcers in executive monkeys. *Sci. Am.* **199**, 95–100.

COCHRANE, R. (1969). Neuroticism and the discovery of high blood pressure. *J. psychosom. Res.* **13**, 21–25.

CRUZ-COKE, R. (1960). Environmental influences and arterial blood pressure. *Lancet* ii, 885–886.

DYKMAN, R. A. & GANTT, W. H. (1960). Experimental psychogenic hypertension: blood pressure changes conditioned to painful stimuli. *Bull. Johns Hopkins Hosp.* **107**, 72–89.

ENGEL, B. T. & BICKFORD, A. F. (1961). Response specificity, stimulus response and individual response specificity in essential hypertensives. *Archs gen. Psychiat.* **5**, 478–489.

FARRIS, E. J., YEAKEL, E. H. & MEDOFF, H. S. (1945). Development of hypertension in emotional grey Norway rats after air blasting. *Am. J. Physiol.* **144**, 331–333.

GAMBARO, S. & RABIN, A. I. (1969). Diastolic blood pressure responses following direct and displaced aggression after anger arousal in high and low guilt subjects. *J. Pers. soc. Psychol.* **12**, 87–94.

GEIGER, H. J. & SCOTCH, N. A. (1963). The epidemiology of essential hypertension: a review with special attention to psychologic and sociocultural factors. I. Biologic mechanisms and descriptive epidemiology. *J. chron. Dis.* **16**, 1151–1182.

GENEST, J. *et al.* (1961). Studies of the pathogenesis of human hypertension: the adrenal cortex and renal pressor mechanism. *Ann. int. Med.* **55**, 12–28.

GHOSH, B. N. & BASHEY, H. I. (1966). An exploratory study of psychological and social factors in essential hypertension amongst Indian patients. *J. Indian med. Ass.* **47**, 267–272.

GLOCK, C. Y. & LENNARD, M. L. (1957). Studies in hypertension. V. Psychologic factors in hypertension: an interpretative review. *J. chron. Dis.* **5**, 174–185.

GOLDBERGER, E. (1958). The pathogenesis of essential hypertension: a unified theory. *Am. J. Cardiol.* **1**, 154–175.

GRAHAM, J. D. P. (1945). High blood pressure after battle. *Lancet* i, 239–240.

GRESSEL, G. C. *et al.* (1949). Personality factors in arterial hypertension. *J. Am. med. Ass.* **140**, 265–272.

HARBURG, E., MCGINN, N. F. & WIGLE, J. B. (1965). Recalled treatment by parents among college males and blood pressure level *vs.* variability. *J. psychosom. Res.* **9**, 173–183.

HARBURG, E. *et al.* (1964). Personality traits and behavioural patterns associated with systolic blood pressure levels in college males. *J. chron. Dis.* **17**, 405–414.

HARRIS, R. E. (1953). Dominance assertiveness and hostility in persons with high and low blood pressure: some similarities and some differences. (Prepublication research memorandum.)

HINKLE, L. E. (1961). Ecological observations of the relation of physical illness, mental illness and the social environment. *Psychosom. Med.* **23**, 289–297.

HINKLE, L. E. & WOLFF, H. G. (1961). The role of emotional and environmental factors in essential hypertension. (Prague symposium.)

HOKANSON, J. E. & BURGESS, M. (1962*a*). The effects of three types of aggression on vascular processes. *J. abnorm. soc. Psychol.* **64**, 446–449.

HOKANSON, J. E. & BURGESS, M. (1962*b*). The effects of status, type of frustration, and aggression on vascular processes. *J. abnorm. soc. Psychol.* **65**, 232–237.

HOKANSON, J. E., BURGESS, M. & COHEN, M. F. (1963). Effects of displaced aggression on systolic blood pressure. *J. abnorm. soc. Psychol.* **67**, 214–218.

HOKANSON, J. E. & SHETLER, S. (1961). The effect of overt aggression on physiological arousal level. *J. abnorm. soc. Psychol.* **63**, 446–448.

HOLMES, D. S. (1966). Effects of overt aggression on level of physiological arousal. *J. Person. soc. Psychol.* **4**, 189–194.

Hudak, W. J. & Buckley, J. P. (1961). Production of hypertensive rats by experimental stress. *J. pharm. Sci.* **50**, 263–264.

Kalis, B. L. *et al.* (1961). Personality and life history factors in persons who are potentially hypertensive. *J. nerv. ment. Dis.* **132**, 457–468.

Kuntz, A. (1953). *The Autonomic Nervous System.* Philadelphia: Saunders.

Lacey, J. I. (1950). Individual differences in somatic response patterns. *J. comp. physiol. Psychol.* **43**, 338–350.

Lacey, J. I. (1956). The evaluation of autonomic responses: towards a general solution. *Ann. N.Y. Acad. Sci.* **67**, 125–163.

Lacey, J. I., Bateman, D. E. & Van Lehn, R. (1953). Autonomic response specificity: an experimental study. *Psychosom. Med.* **15**, 8–21.

Lacey, J. I. & Lacey, B. C. (1958). The relationship of resting autonomic activity to motor impulsivity. In *The Brain and Human Behaviour.* Baltimore: Williams & Wilkins.

Lacey, J. I. & Smith, R. L. (1954). Conditioning and generalization of unconscious anxiety. *Science, N.Y.* **120**, 1045–1052.

Lacey, J. I., Smith, R. L. & Green, A. (1955). Use of conditioned autonomic responses in the study of anxiety. *Psychosom. Med.* **17**, 208–217.

Ledingham, J. M. (1951). The nature of hypertension occurrring in the nephrectomized parabiotic rat. *Clin. Sci.* **10**, 423.

McGinn, N. F. *et al.* (1964). Psychological correlates of blood pressure. *Psychol. Bull.* **61**, 209.

Maher, B. A. (1966). *Principles of Psychopathology.* New York. McGraw-Hill.

Malmo, R. B. (1957). Anxiety and behavioural arousal. *Psychol. Rev.* **64**, 276–287.

Medoff, H. S. & Bongiovanni, A. M. (1945). Blood-pressure in rats subjected to audiogenic stimulation. *Am. J. Physiol.* **143**, 300–305.

Mendlowitz, M. *et al.* (1962). Vascular reactivity in the patient with essential hypertension and hypertension of renal origin. *Am. J. Cardiol.* **9**, 680.

Mills, L. C. (1959). Hypertension and stress. In J. Moyer (ed.), *Hypertension*—First Hahnemann Symposium on Hypertensive Disease. Philadelphia: Saunders.

Morris, J. N. *et al.* (1953). Coronary heart disease and physical activity of work. *Lancet* ii, 1053–1057; 1111–1120.

Morrison, S. L. & Morris, J. N. (1959). Epidemiological observations on high blood pressure without apparent cause. *Lancet* ii, 864–870.

Robinson, J. O. (1963). A study of neuroticism and casual arterial blood pressure. *Br. J. soc. clin. Psychol.* **2**, 56–64.

Ruskin, A., Beard, O. W. & Schaffer, R. L. (1948). Blast hypertension: elevated arterial pressure in victims of the Texas City disaster. *Am. J. Med.* **4**, 228.

Sainsbury, P. (1960). Neuroticism in unselected out-patients attending physical medicine and orthopaedic departments. *Ann. physical Med.* **5**, 310–317.

Sawrey, W. L. & Weisz, J. D. (1956). An experimental method of producing gastric ulcers. *J. comp. physiol. Psychol.* **49**, 269–270.

Scotch, H. A. & Geiger, H. J. (1963). The epidemiology of essential hypertension: a review with special attention to the psychologic and sociocultural factors. II. Psychologic and sociocultural factors in etiology. *J. chron. Dis.* **16**, 1183–1213.

Shapiro, A. P. (1960). Psychophysiologic mechanisms in hypertensive vascular disease. *Ann. int. Med.* **53**, 64–83.

Shapiro, A. P. & Melhado, J. (1958). Observations on blood pressure and other physiologic and biochemical mechanisms in rats with behavioural disturbances. *Psychosom. Med.* **20**, 303–313.

Simonson, E. & Brozek, J. (1959). Russian research on arterial hypertension. *Ann. int. Med.* **50**, 129–193.

Sines, J. O. (1962). Strain differences in activity, emotionality, body weight and susceptibility to stress induced stomach lesions. *J. genet. Psychol.* **101**, 209–217.

Stevenson, I. P. *et al.* (1952). Hypertension as a reaction pattern to stress. *Am. J. med. Sci.* **224**, 286–299.

Stokvis, B. (1953). The problem of specificity in psychosomatic practice investigated by structural-analytical means. *Ned. Tijdschr. Geneek.* **97**, 30–43.

Van Der Valk, J. M. (1957). Blood pressure changes under emotional influences in patients with essential hypertension and control subjects. *J. psychosom. Res.* **2**, 134–146.

WILHELMJ, C. M. *et al.* (1953). Emotional elevations of blood pressure in trained dogs: possible relation to origin of hypertension in humans. *Psychosom. Med.* **15**, 390–395.

WOLF, S. *et al.* (1948). Hypertension as a reaction pattern to stress. *Ann. int. Med.* **29**, 1056–1076.

WOLF, S. *et al.* (1955). *Life Stress Essential Hypertension: a Study of Cardiovascular Adjustments in Man.* Baltimore: Williams & Wilkins.

Manuscript received 25 *September* 1969

Revised manuscript received 19 *February* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 73–78
With 1 text-figure
Printed in Great Britain

Smoking Behaviour and its Relation to the Smoker's Immediate Experience

By C. D. FRITH

Institute of Psychiatry, London University

Although there have been many large surveys of smoking behaviour, as yet there is little data that can help to answer questions like 'Why do people smoke?' and more particularly 'Why do people smoke tobacco?' What evidence there is suggests that the vital component of the tobacco for most smokers is nicotine (Lucchesi *et al.*, 1967). Hence it seems plausible to suppose that people smoke, at least in part, because of the psychopharmacological effects that nicotine has upon them. The pharmocological effects of nicotine have been quite widely investigated, particularly in animals. In most situations it has been found that nicotine acts as a stimulant, increasing the level of 'arousal' of the organism (Armitage *et al.*, 1969). The various levels of arousal are conceived as forming a continuum of mental activity, varying from deep, dreamless sleep to extreme excitement. In certain cases nicotine appears to decrease arousal. For example, when nicotine was given intravenously to cats in a dose of 2 μg/kg every 30 sec. for 20 min., it caused desynchronization of the electrocorticogram, indicating cortical activation, and an increase in the release of cortical acetylcholine. However, a larger dose given less frequently (4 μg/kg every minute for 20 min.) caused sometimes an increase and sometimes a decrease in cortical activity, such changes being accompanied by an increase or a decrease in cortical acetylcholine output (Armitage *et al.*, 1969). Thus it can be proposed that people use the nicotine in cigarettes to produce either an increase or a decrease in their level of arousal (Burn, 1968).

If nicotine acts as a stimulant, then a person is most likely to smoke when he has an undesirably low level of arousal. Such a low level of arousal would probably result from a combination of circumstances. The person would be particularly prone to low levels of arousal by his very nature and would also be in a situation inducive of a particularly low level of arousal (tiring, boring, etc.). Another person might smoke because he was particularly prone to high levels of arousal and found himself in a highly arousing situation. It seems very likely that these two would be different people. Thus it should be possible to isolate two extremes of smoking behaviour. There should be one group of people who smoke mostly in situations inducing low levels of arousal in order to increase their arousal level, and there should be another group of people who smoke in situations which induce high levels of arousal in order to reduce their arousal level. What follows concerns the first of a series of experiments designed to test this hypothesis.

A questionnaire was constructed in which situations inducive of high and low arousal were described. The subjects had to imagine themselves in each of these situations and indicate what their desire for a cigarette would be on a seven-point scale. There were 12 high-arousal situations and 10 low-arousal situations. The 12 high-arousal situations concerned emotional stress and anxiety (e.g. you are watching a violent and horrifying film), of which there were six samples (1, 5, 9, 12, 16, 20) and also stress induced by mental activity (e.g. you have to do some rapid mental arithmetic for an intelligence test), of which there were six samples (2, 6, 10, 13, 17, 21). The 10 low-arousal situations concerned relaxation and boredom (e.g. you are having a restful evening alone reading a magazine), of which there were six samples (4, 7, 8, 15, 18, 19) and also bodily tiredness and repetitive work (e.g. you are very tired and need to keep awake), of which there were four samples (3, 11, 14, 22). All these questions and the instructions for answering them are given in the appendix.

There were 98 subjects (59 men, 39 women), all employees of the Institute of Psychiatry, and their friends. They represented a wide range of occupations and were aged between 18 and 50. All were cigarette smokers. Table 1 shows the means and standard deviations of level of desire for a cigarette in each of the 22 situations, for the number of cigarettes smoked per day and for age. It can be seen that low-arousal situations tended to induce more desire for a cigarette than high-arousal situations. Table 2 shows the intercorrelations between the 25 variables (the additional one being sex). Most of the situations are positively correlated, usually significantly so. This is largely due to the variable concerning the number of cigarettes smoked per day, heavy smokers having a high desire for a cigarette in any situation. Fig. 1 shows a graphical representation of the first two factors extracted from the correlation matrix by a principal component analysis. These first two factors together accounted for 31·3 per cent of the variance. On the first unrotated factor the number of cigarettes smoked per day loads most highly with all the situations loading in the same direction. The second factor has high-arousal situations loading in one direction and low-arousal situations loading in the other direction. The existence of this factor provides some support for the hypothesis that there are some people who smoke most when in high-arousal situations and other people who smoke most when in low-arousal situations. An unexpected result was that sex loaded very highly on this factor, indicating that the men wished to smoke in low-arousal situations, whereas the women wished to smoke in high-arousal situations. Since many of the requirements for the factor-analytic technique used might not have been fulfilled in these data (e.g. the ratings of desire for a cigarette might form a ordinal scale only), a further check was made by comparing the desire of men and women for cigarettes in various situations using *t*-tests. Table 3 shows these tests for the most extreme situations indicated by the principal component analysis. The *t*-tests confirmed that men had a higher desire for a cigarette in the low-arousal situations, while women had a higher desire in the high-arousal situations.

Table 1. *Means and S.D.s for age, cigarettes per day and desire for a cigarette induced by the various situations*

(The situations are arranged in order of mean desire induced.)

Source	Mean	S.D.	Arousal	Source	Mean	S.D.	Arousal
Age (years)	32·8	12·5					
Cigs./day	18·5	8·5					
Q. 18	3·4	1·5	L	1	2·2	1·9	H
4	3·3	1·4	L	13	1·9	1·5	H
22	2·9	1·6	L	2	1·9	1·7	H
15	2·8	1·4	L	5	1·9	1·8	H
11	2·7	1·7	L	12	1·9	1·9	H
7	2·7	1·4	L	3	1·6	1·6	L
20	2·5	1·8	H	19	1·5	1·6	L
8	2·4	1·5	L	14	1·3	1·3	L
16	2·4	2·0	H	21	1·2	1·7	H
9	2·4	2·1	H	17	1·2	1·8	H
10	2·3	1·6	H	6	0·7	1·5	H

In general, it appeared that people had a greater desire to smoke in low-arousal situations. This is consistent with nicotine acting as a stimulant. The finding that heavy smokers have a desire for a cigarette in any situation was not unexpected, but is of importance since this effect will tend to disguise any differential patterns of smoking behaviour that may exist. When the data were examined in more detail, evidence appeared for the existence of two extremes of smoking behaviour. The men tended to have the highest desire for a cigarette in situations inducing boredom and tiredness, while the women had their highest desire in stress inducing situations. McKennell (1968) attempted to type smokers by occasions for smoking and found two factors which he called 'nervous irritation smoking' and 'relaxation smoking' which are probably very similar to those found here. However, he apparently did not relate these factors to sex. The relationship of smoking behaviour to sex was unexpected, but receives partial support from an American survey on the use of tobacco (Horn, 1969). As a small part of this survey (question 252) people were asked how much they agreed with statements like 'when I feel uncomfortable or upset about something, I light up a cigarette'.

There were five such statements with which women agreed significantly more often than men. All described situations of stress and high arousal (e.g. feeling worried, angry, upset, etc.). Unfortunately there were no questions describing low-arousal states. However, with one statement, 'I smoke cigarettes to stimulate me', more men than women agreed, although not significantly so.

Table 2. *Correlation matrix for sex, cigarettes per day, age and the 22 situations*

($n=98$; $r=27$ at the 0·01 level of significance.)

Sex	—																								
Cigs./day	10	—																							
Age	−40	−06	—																						
Q. 1	03	22	−30	—																					
2	15	19	00	22	—																				
3	03	11	−10	09	26	—																			
4	−24	25	09	−16	13	21	—																		
5	−15	30	−01	11	06	12	25	—																	
6	−10	33	03	38	08	04	00	22																	
7	30	21	−26	11	31	12	06	02	−08																
8	11	32	−13	00	31	39	45	46	07	31															
9	04	30	−17	54	32	18	−02	12	19	22	00	—													
10	−30	25	−05	10	11	24	24	30	05	05	32	04	—												
11	11	29	−17	09	28	14	12	−05	09	41	21	14	23	—											
12	−02	25	−25	62	18	02	−04	05	37	10	03	38	21	13											
13	13	37	−09	19	31	33	22	25	32	12	42	16	16	09	29	—									
14	03	16	−06	18	27	48	19	23	22	27	46	03	20	18	13	44	—								
15	21	31	−21	22	21	−01	20	13	10	43	34	08	17	27	12	07	29	—							
16	−21	16	09	33	10	03	−02	29	35	−04	09	24	10	−15	24	14	25	−02	—						
17	−08	25	04	44	15	17	11	18	39	09	−01	35	09	18	40	27	22	01	36	—					
18	00	22	−09	17	06	18	40	21	−04	33	41	−04	14	24	−09	15	33	40	07	−01	—				
19	16	35	04	17	12	12	24	23	05	28	41	24	06	−02	15	24	36	21	13	26	17	—			
20	−12	15	−15	37	11	01	12	21	29	00	06	30	18	−09	30	18	17	06	57	32	02	25	—		
21	07	33	−13	35	13	07	11	18	45	10	19	13	20	25	27	36	29	32	19	45	01	15	32	—	
22	29	28	−23	25	21	04	06	−07	06	26	23	20	03	38	21	12	09	45	−23	−02	02	13	−05	29	—

The original hypothesis of this experiment was that smoking behaviour would be determined by the smoker's characteristic level of arousal. On this basis the results suggest that women are usually more highly aroused than men. While it is true that women are more neurotic than men (Eysenck & Eysenck, 1964), there is an alternative explanation of the results. In a sample of over 4000 cigarette smokers, Todd (1969) found that 77·7 per cent of the men inhaled the smoke whereas only 53·0 per cent of the women did so.

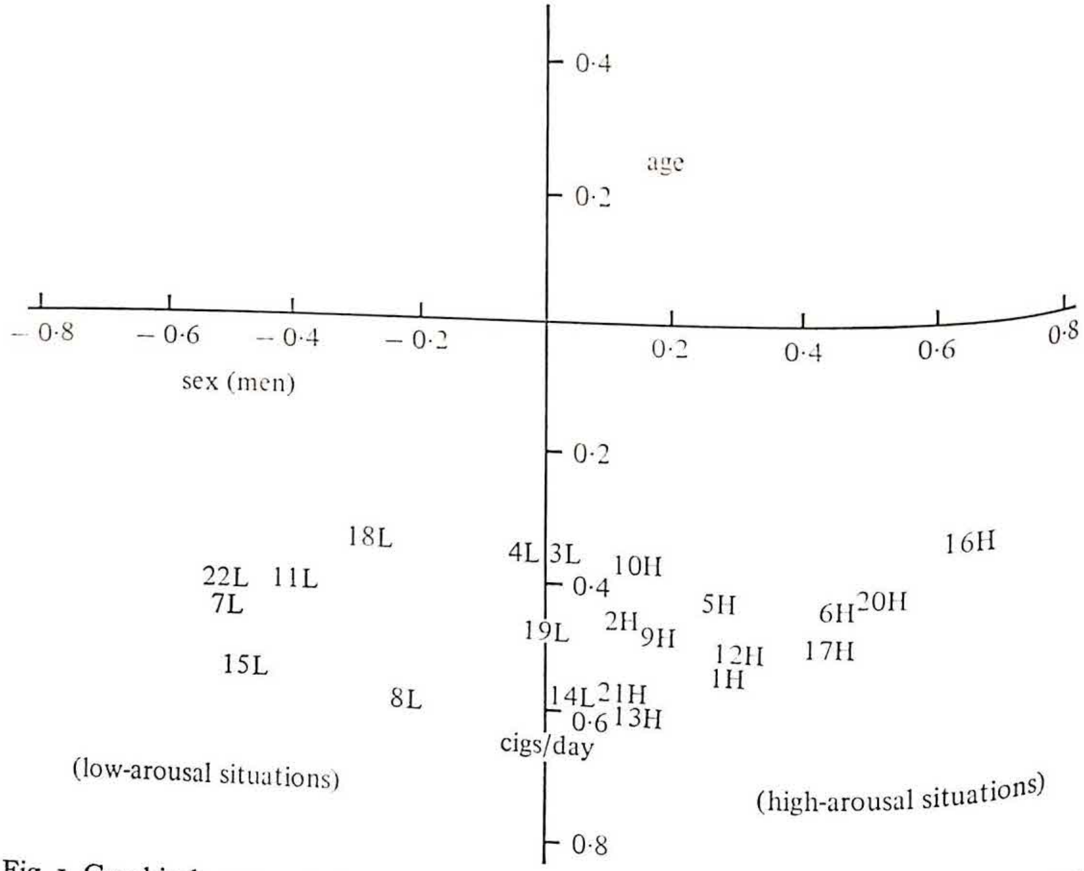


Fig. 1. Graphical representation of the 22 situations, age, sex and cigarettes per day showing their loadings on the first two factors of a principal components analysis.

Table 3. *A comparison of the desire for a cigarette in men and women for the six most extreme situations*

Question	Men		Women			
	Mean	S.D.	Mean	S.D.	*t*	*P*
			High Arousal			
16	2·05	1·82	2·87	2·08	2·06	< 0·05
20	2·32	1·62	2·77	2·06	1·21	n.s.
6	0·56	1·30	0·87	1·76	1·00	n.s.
			Low arousal			
22	3·29	1·47	2·33	1·64	3·02	< 0·01
7	3·02	1·29	2·18	1·30	3·15	< 0·01
15	3·00	1·27	2·38	1·55	2·17	< 0·05

People who do not inhale absorb virtually no nicotine (Kuhn, 1965) and thus the pharmacological effects of nicotine would be irrelevant to their smoking behaviour. Therefore if the people who reported wishing to smoke under stress tended not to inhale their cigarettes, then their behaviour requires an explanation quite different from that originally proposed. For example, the ritual of lighting and holding a cigarette might be important in reducing the effect of the stressful situation.

This work was supported by a grant from the Tobacco Research Council.

REFERENCES

ARMITAGE, A. K., HALL, G. H. & SELLERS, C. M. (1969). Effects of nicotine on electrocortical activity and acetylcholine release from the cat cerebral cortex. *Br. J. Pharmacol.* **35**, 152–160.

BURN, J. H. (1968). Letter. *Br. med. J.* (3 Feb.), 310–311.

EYSENCK, H. J. & EYSENCK, S. B. G. (1964). *Manual of the Eysenck Personality Inventory.* London: University of London Press.

HORN, D. (1969). *Use of Tobacco.* U.S. Department of Health, Education and Welfare. National Clearing House for Smoking and Health.

KUHN, H. (1965). Tobacco alkaloids and their pyrolisis products in the smoke. In V. Euler (ed.), *Tobacco Alkaloids and Related Compounds.* Oxford: Pergamon Press.

LUCCHESI, B. R., SCHUSTER, C. R. & EMPLEY, G. S. (1967). The role of nicotine as a determinant of cigarette smoking frequency in man with observations of certain cardiovascular effects associated with the tobacco alkaloid. *Clin. Pharmacol. Therapeut.* **8**, 787–796.

MCKENNELL, A. C. (1968). British research into smoking behavior. In E. F. Borgatta & R. R. Evans (eds.), *Smoking, Health and Behavior.* Chicago: Aldine.

TODD, G. F. (ed.) (1969). *Statistics of Smoking in the United Kingdom.* London: Tobacco Research Council.

Manuscript received 16 February 1970

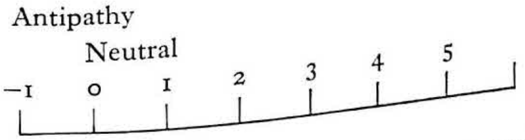
Revised manuscript received 20 April 1970

APPENDIX

Situational Smoking Questionnaire

Age____________________ Sex____________________

In the questions that follow you are to imagine yourself in the situations described and indicate what you think your craving for a cigarette would be in each of the situations. The degree of craving should be indicated in terms of the scale shown here:



Thus for each situation you must write down the number corresponding to what you think your desire for a cigarette would be. Please remember that we are interested only in your desire for a cigarette in the various situations and not whether it would be proper or possible to smoke in these situations.

1. Your are having an important interview for a job.
2. You have to fill in a complicated tax form.
3. You have to look through several hundred coins to see if there are any rare and valuable ones.
4. You are having a quiet evening with friends.
5. You are witnessing a violent and horrifying film.
6. You have to drive at speed in heavy traffic.
7. You have to wait for your train home, which is very late.

8. You are having a restful evening alone reading a magazine.
9. You are sitting in a dentist's waiting room knowing that you are to have a particularly difficult filling.
10. You are trying to hold a conversation at a large and very noisy party.
11. You are very tired and need to keep awake.
12. You have to ask your boss for a rise at a time when he is known to be in a bad mood.
13. You are trying to account for the discrepancy between your spending for the month and your bank statement.
14. You are looking through a long list of names to see if you can find anyone you know.
15. You are chatting with friends during a tea-break.
16. You have just been informed of the death of a close friend.
17. You have to do some rapid mental arithmetic for an intelligence test.
18. You are travelling on a train for several hours.
19. You go for a solitary walk in quiet countryside.
20. You have just heard the announcement of a plane crash and you think a friend may have been involved.
21. You are having an important telephone conversation in a very noisy office.
22. You have just had a very big meal.

Br. J. soc. clin. Psychol. (1971), **10**, pp. 79–80
Printed in Great Britain

'A New Measure of Conservatism': its Limitations

By JOHN J. RAY

Macquarie University, Australia

Wilson & Patterson (1968) have presented a conservatism scale having items in a new short format. This format is proposed as intrinsically more reliable and more valid. It is also said to be less susceptible to acquiescent response set.

A defect in its standardization is that none of the samples to which it was administered appeared to approximate a normal one. Most of the groups named appeared to have been contacted through educational institutions.

In Australia the conscription laws provide for a random selection of recruits by a ballot procedure. The resulting sample is nationwide and hence ideal for social research. It is, however, limited to male 20-year-olds. To this sample a questionnaire was administered by uniformed Army psychologists as the 'research project of a student at the University of Sydney'. It included the C-scale. Two hundred and seventy-six completed protocols were received from the October 1969 intake at Kapooka training battalion.

The questionnaire was identical to that presented by Wilson & Patterson (1968) except that the item 'Birching' had been changed to 'Flogging'. All odd-numbered items were reverse scored. During scoring a very notable feature was the great preponderance of 'yes' responses.

The reliability observed (coefficient 'alpha') was 0·63. This compares with 0·94 (split-half) reported by Wilson & Patterson (1968).

In a second check of the C-scale with 262 first-year psychology students (a sample more comparable to that of Wilson & Patterson) reliability was a reasonably satisfactory 0·83. The mean and standard deviation were 112·76 and 13·17 respectively.

It is concluded that the C-scale is not suitable for use with a sample randomly drawn from the general population. The lack of discriminating power of the individual items that can be noted from Table 1 is probably due to the low variability inherent in a strong acquiescent response set.

The item means of Table 1 are based on a scoring system of 3, 2, 1, instead of the original 2, 1, 0. It is normal practice in coding data on to computer cards to reserve zero for representing 'no response'. As needed, zeros can be transformed in the computer to the midpoint of the scale (as was done here).

Table 1. *Item and scale total statistics for 276 Australian National Servicemen*

(All item-total correlations are corrected for spurious overlap. A high mean indicates low conservatism.)

Item	Mean	S.D.	Item-total correlation
Death penalty	2·36	0·48	0·059
Evolution theory	1·55	0·50	0·132
School uniforms	2·78	0·42	0·286
Striptease shows	1·30	0·46	0·262
Sabbath observance	2·44	0·50	0·199
Beatniks	1·66	0·47	0·402
Patriotism	2·57	0·49	0·027
Modern art	1·36	0·48	0·170
Self-denial	2·46	0·50	0·165

Table 1—*continued*

Item	Mean	S.D.	Item-total correlation
Working mothers	1·65	0·48	−0·026
Horoscopes	2·19	0·39	−0·094
Birth control	1·20	0·40	0·111
Military drill	2·55	0·50	0·239
Co-education	1·11	0·32	0·146
Divine law	2·42	0·49	0·290
Socialism	1·78	0·41	−0·017
White superiority	2·07	0·25	0·033
Cousin marriage	1·66	0·48	0·223
Moral training	2·82	0·39	0·163
Suicide	1·94	0·23	0·117
Chaperones	2·21	0·40	0·159
Legalized abortion	1·34	0·48	0·365
Empire building	2·39	0·49	0·045
Student pranks	1·64	0·48	0·240
Licensing laws	2·70	0·46	0·027
Computer music	1·64	0·48	0·112
Chastity	2·38	0·49	0·199
Fluoridation	1·22	0·41	−0·037
Royalty	2·47	0·50	0·140
Women judges	1·35	0·48	−0·108
Conventional clothes	2·67	0·47	0·102
Teenage drivers	1·13	0·33	0·134
Apartheid	2·08	0·28	0·061
Nudist camps	1·37	0·48	0·379
Church authority	2·36	0·48	0·317
Disarmament	1·57	0·50	0·060
Censorship	2·14	0·49	0·309
White lies	1·55	0·50	0·120
Flogging	2·17	0·37	0·070
Mixed marriage	1·26	0·44	0·174
Strict rules	2·36	0·48	0·197
Jazz	1·25	0·43	0·089
Straitjackets	2·32	0·47	0·002
Casual living	1·16	0·37	0·188
Learning Latin	2·28	0·45	0·113
Divorce	1·50	0·50	0·304
Inborn conscience	2·47	0·50	0·004
Coloured immigration	1·43	0·50	0·014
Bible truth	2·52	0·50	0·338
Pyjama parties	1·21	0·41	0·241
Total scale	96·30	5·17	

REFERENCE

WILSON, G. D. & PATTERSON, J. R. (1968). A new measure of conservatism. *Br. J. soc. clin. Psychol.* 7, 264–269.

Manuscript received 13 March 1970

Revised manuscript received 31 June 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 81–83
With 1 text-figure
Printed in Great Britain

A Note on the Validity of the Wilson–Patterson Measure of Conservatism

BY JAMES HARTLEY AND JANET HOLT

University of Keele

Wilson & Patterson (1968) published a new measure of conservatism consisting of a list of 50 brief labels or catchphrases representing various familiar and controversial issues to which the respondent has to indicate whether or not he favours or believes in them by circling as appropriate 'yes', 'no' or '?'.

During the last year or so we have been trying to develop a scale to measure teachers' attitudes to new educational media (Hartley & Holt, 1970), and in the course of this investigation we have often, in addition, administered the Wilson–Patterson scale. The first part of this note concerns itself with some of the data that we have collected, and the second with some observations on the technique used.

In collecting our data we have used only the first half of the scale (Wilson & Patterson report a correlation of 0·943 with $n = 244$ between the two halves). The means and standard deviations of the subgroups to whom we have administered the scale are shown in Table 1 in increasing order of conservatism.

Table 1. *Means and S.D.s obtained on half of the Wilson–Patterson Conservatism Scale*

Group	n	Mean	S.D.
Second-year psychology undergraduate students*			
Male	32	12·9	5·7
Female	33	14·3	6·0
Teacher counsellors			
Male	11	14·4	6·5
Female	7	16·6	7·6
Fourth-year education students*			
Male	7	13·1	7·9
Female	20	19·3	8·0
Postgraduate education students			
Male	21	17·3	9·1
Female	8	17·1	5·7
Teachers			
Male	21	17·8	4·2
Female	13	18·2	6·9
Deputy heads			
Male	16	19·3	8·4
Female	12	15·3	6·2
College of education students			
Male	18	19·4	8·2
Female	34	22·7	6·0
Headmasters			
Male	15	19·9	6·1

* Keele University provides a four-year course: after the first general year, students take two principal and two subsidiary subjects. Education may be one of these principal subjects.

What we feel is important to note here is (1) these norms appear to be (as expected) lower than New Zealand counterparts quoted by Wilson & Patterson; (2) that the second-year undergraduate students are significantly less conservative than their fellow college of education students of the same age ($P < 0.01$); and (3) that teacher counsellors might be less conservative than 'teachers'.

Together with the fact that the Wilson–Patterson scale is remarkably quick and easy to use, and is popular with the respondents, these observations lead us to the conclusion that the scale is one of considerable value. The pattern of results shown in Table 1, for instance, could profitably be explored further.

Wilson & Patterson argue that traditional attitude scales, which present items in the form of propositional statements to which the respondent must record some level of agreement or disagreement, 'can never provide an adequate basis for the measurement of attitudes'. Amongst the many criticisms they voice, we may note (i) 'the danger of acquiescence response bias is enhanced because the propositions are directive and complex, often requiring multi-negative thinking'; (ii) 'catches are suspected, and critical reaction is adjusted in the direction of "social desirability"'; and (iii) 'the eventual response becomes more a function of cognitive processes than affective'. They see their scale as 'an improvement insofar as it reduces the influences of cognitive processes, task conflict, grammatical confusion and social desirability'.

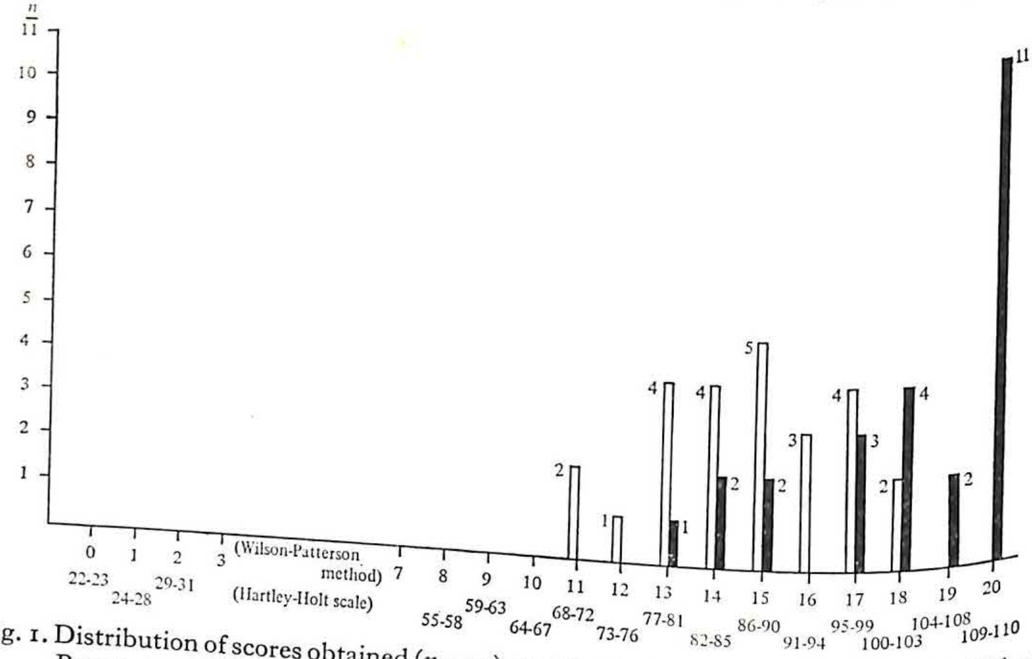


Fig. 1. Distribution of scores obtained ($n = 25$) on attitudes to new media, using (i) the Wilson–Patterson method (■), and (ii) the Hartley–Holt attitude scale (□).

With most of these statements we would agree: they apply forcefully to our own scale, designed to measure teachers' attitudes to new educational media, which has been developed using standard Likert procedures. We are not so sure, however, about the last point. Teachers (to us) appear to be prone to make socially desirable responses in attitude questionnaires. For example, the mean score of 18 teacher counsellors completing our attitude questionnaire was 88·7 (S.D. 6·3): when asked to complete it again as 'a good teacher' might, the results were a mean score of 90·6 (S.D. 9·7).

To investigate these issues we did two things. First we correlated the scores of the college of education students on the Wilson–Patterson scale with their scores on the Crowne–Marlowe (1960) scale of social desirability. The results were: for men ($n = 18$), $r = +0.15$; for women ($n = 34$), $r = +0.38$ ($P < 0.05$); and for the total sample ($n = 52$), $r = +0.31$ ($P < 0.05$).

Next we inserted at regular intervals 10 new items into the Wilson–Patterson scale, items representing controversial issues in education, and we asked a group of 25 teachers to complete this new version of the scale, and our own teacher-attitude questionnaire. The results of this investigation are shown graphically in Fig. 1. It is clear that both scales produced 'favourable' responses, but that there was more discrimination between teachers with our questionnaire. (The correlation between the two scales was in fact −0·006.) Table 2 shows the distribution of scores obtained from 34 teachers on each of the extra 10 items, and these results reflect the findings shown in Fig. 1.

Table 2. *The distribution of teachers indicating that 'they favoured or believed in' the following items inserted into the Wilson–Patterson Conservatism Questionnaire* ($n = 34$)

Item	Yes	?	No
Educational television	33	1	0
Teaching machines	26	7	1
Language laboratories	28	6	1
Programmed instruction	27	7	0
Audio-visual aids	34	0	0
Computer-assisted instruction	22	12	0
Classroom radio	29	4	1
Educational films	33	1	0
Programmed texts	28	3	3
Educational wallcharts	27	5	2

It seems clear from these two different approaches to the problem that 'social desirability' can affect questionnaires employing the techniques advocated by Wilson & Patterson.

REFERENCES

Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *J. consult. Psychol.* **24**, 349–354.

Hartley, J. & Holt, J. (1970). Measuring teachers' attitudes to new educational media. (Available from the authors.)

Wilson, G. D. & Patterson, J. R. (1968). A new measure of conservatism. *Br. J. soc. clin. Psychol.* **7**, 264–269.

Manuscript received 16 March 1970

Revised manuscript received 30 April 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 84–86
With 1 text-figure
Printed in Great Britain

Measuring Social Attitudes in Children

By PAUL INSEL and GLENN D. WILSON

Institute of Psychiatry, London

The Conservatism Scale (Wilson & Patterson, 1968) was developed on the assumption that an important general factor underlies all social attitudes, including those related to religion, politics, art, science, law, punishment and sexual morality. This hypothesis, originally supported with a New Zealand sample, was later confirmed in England (Wilson, 1970) and the Netherlands (Bagley *et al.*, 1970). In each case very high internal consistency was reported (split-half coefficients of 0·94, 0·93 and 0·89 respectively). The usefulness and construct validity of the scale have been demonstrated by relating conservatism scores to several diverse areas of behaviour: club membership (Wilson & Patterson, 1968), humour preferences (Wilson & Patterson, 1969), conditioning (Wilson, 1968), self-concept (Boshier, 1969), self-identification (Patterson & Wilson, 1969) and racial prejudice (Bagley, 1970).

The outstanding feature of the C scale is the simplicity of its item format, which makes for exceptionally easy administration and scoring. Directional statement-form items used in traditional attitude scales are replaced by short catch-phrases representing familiar social issues, and the respondent is required to indicate his position on a three-point scale.

Table 1. *The Children's Conservatism Scale: Instructions, items and format*

Which of the following do you favour or believe in ?

(Circle 'Yes' or 'No'. If absolutely uncertain, circle '?'. There are no right or wrong answers; do not discuss; just give your first reaction.)

1. Hanging thieves	Yes	?	No	26. Computers	Yes	?	No
2. Space travel	Yes	?	No	27. Saying prayers	Yes	?	No
3. School uniforms	Yes	?	No	28. Going barefoot	Yes	?	No
4. Bikinis	Yes	?	No	29. The Queen	Yes	?	No
5. Sunday school	Yes	?	No	30. Women doctors	Yes	?	No
6. Bearded men	Yes	?	No	31. Hard work	Yes	?	No
7. Saluting the flag	Yes	?	No	32. Drinking beer	Yes	?	No
8. Modern art	Yes	?	No	33. Atomic bombs	Yes	?	No
9. Obedience	Yes	?	No	34. Nude swimming	Yes	?	No
10. Comics	Yes	?	No	35. Church	Yes	?	No
11. Miracles	Yes	?	No	36. Chinese food	Yes	?	No
12. Dancing	Yes	?	No	37. Politeness	Yes	?	No
13. Army drill	Yes	?	No	38. Telling fibs	Yes	?	No
14. Mixed schools	Yes	?	No	39. Whipping criminals	Yes	?	No
15. The Ten Commandments	Yes	?	No	40. The Germans	Yes	?	No
16. The Russians	Yes	?	No	41. Strict rules	Yes	?	No
17. Whites only	Yes	?	No	42. The Beatles	Yes	?	No
18. Kissing	Yes	?	No	43. Killing enemies	Yes	?	No
19. Strapping bad boys	Yes	?	No	44. Laughing in class	Yes	?	No
20. Swearing	Yes	?	No	45. Fox hunting	Yes	?	No
21. Servants	Yes	?	No	46. Divorce	Yes	?	No
22. Miniskirts	Yes	?	No	47. Confessing sins	Yes	?	No
23. Saving money	Yes	?	No	48. Coloured people	Yes	?	No
24. Playing pranks	Yes	?	No	49. Bible reading	Yes	?	No
25. Policemen	Yes	?	No	50. Playing doctors	Yes	?	

In view of the success of the adult scale it was thought useful to attempt the construction of a children's version, using the original scale as a model.

The original C scale was regarded as suitable for persons over the age of 14. The children's version to be described was designed for use with children aged 8–13. In general the procedure was to translate the broad concepts of the adult scale into more concrete and specific terms, e.g. 'death penalty' became 'hanging thieves', and 'Sabbath observance' became 'Sunday school'. Some items did not require any alteration, however, and certain others had to be replaced altogether. The final set of 50 items (Table 1) was selected following interviews with children of appropriate ages to establish that the meanings were satisfactorily understood.

In order to investigate the concurrent and discriminative validity of the children's C scale, it was administered along with the original scale to a sample of 185 girls aged 12 to 16 (100 from a Catholic convent school, and 85 from a state-supported school), both located in the same middle-class area of south London. Hypotheses were: (1) At this age level the children's and adults' scale should be highly correlated. (2) Both scales should display internal consistency and the ability to discriminate between the two schools.

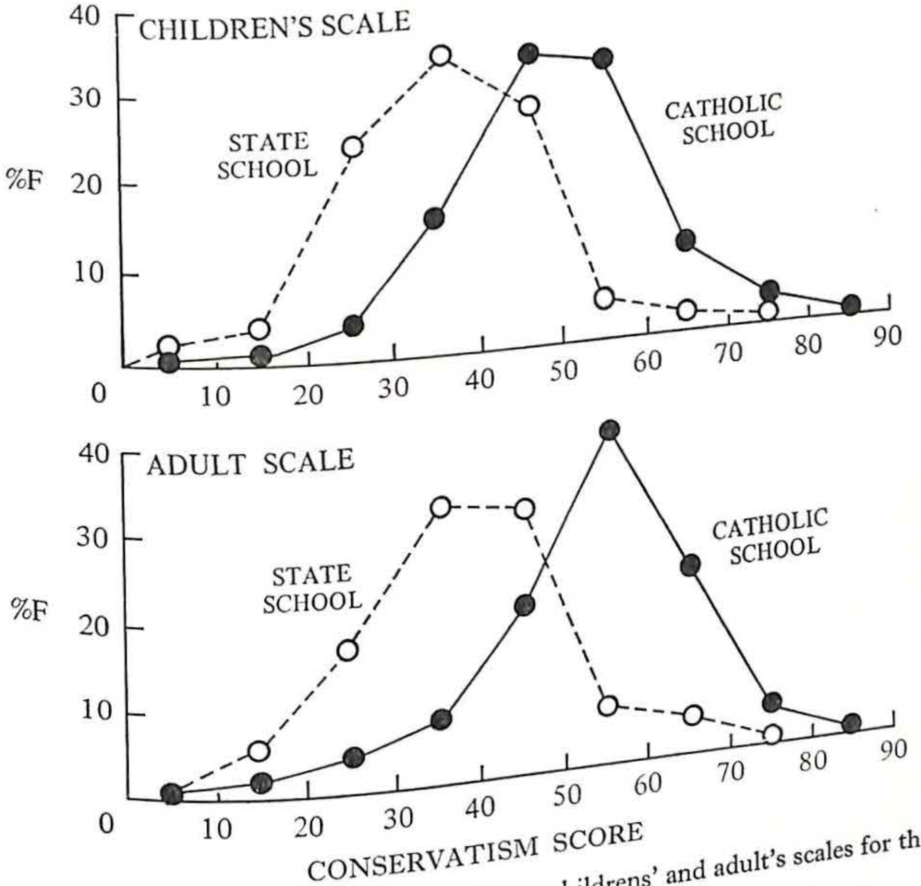


Fig. 1. Distributions of Conservatism scores on the childrens' and adult's scales for the two schools.

Administration and scoring was identical to that reported for the adult scale (Wilson & Patterson, 1968). The scale is entirely self-contained and no scoring key is required because the direction of scoring (conservative or liberal) alternates throughout the test. Each 'yes' to an odd-numbered item scores 2, as does each 'no' to an even-numbered item. Queries and ambiguous responses of any kind score 1. The range of possible scores is therefore 0–100.

The distributions of C scores on the two versions of the test are given for the two schools separately in Fig. 1, and means and S.D.s of these distributions are given in Table 2. Both scales discriminate between the two schools at a high level of significance, although inspection of Fig. 1 suggests that with this age-group the adult scale still has slightly better discriminative validity. The correlation between the adult and children's scales was 0·86 ($n = 185$) which can be regarded as evidence of very satisfactory concurrent validity.

Table 2. *Means and S.D.s of conservatism scores, and the significance of the differences between schools*

	State school ($n = 85$)		Catholic school ($n = 100$)		
	Mean	S.D.	Mean	S.D.	*P* (diff.)
Children's scale	34·64	10·14	47·48	8·77	$< 0·001$
Adult scale	36·86	10·19	51·99	11·30	$< 0·001$

Split-half reliability coefficients were calculated for each scale by correlating total scores on the first 25 items with total scores on the second 25, and applying the Spearman–Brown correction for test length. Both scales showed high internal consistency, but this was slightly higher for the children's scale (0·89 compared to 0·84 for the adult scale). This result may be taken as evidence that the easier items of the children's version have given rise to more systematic responding even with this borderline age-group.

These results indicate that a scale modelled on the Conservatism Scale has considerable promise as a measure of social attitudes in children. Whether the particular scale items used in this study are ideal, especially where other cultures are concerned, is a matter for further study. What does seem clear is that the simplified item format adopted in the present scale does provide a reliable and valid, as well as economical, means of assessing attitudes in school-age children as well as adults.

REFERENCES

BAGLEY, C. (1970). Racial prejudice and the conservative personality. *Pol. Stud.* (in press).

BAGLEY, C., WILSON, G. D. & BOSHIER, R. (1970). The conservatism scale: a factor-structure comparison of English, Dutch and New Zealand samples. *J. soc. Psychol.* (in press).

BOSHIER, R. (1969). A study of the relationship between self-concept and conservatism. *J. soc. Psychol.* **77**, 139–140.

PATTERSON, J. R. & WILSON, G. D. (1969). Anonymity, occupation, and conservatism. *J. soc. Psychol.* **78**, 263–266.

WILSON, G. D. (1968). Personality, GSR conditioning and response to instructional set. *Psychol. Rep.* **22**, 618.

WILSON, G. D. (1970). Is there a general factor in social attitudes? Evidence from a factor analysis of the conservatism scale. *Br. J. soc. clin. Psychol.* **7**, 101–107.

WILSON, G. D. & PATTERSON, J. R. (1968). A new measure of conservatism. *Br. J. soc. clin. Psychol.* **7**, 264–269.

WILSON, G. D. & PATTERSON, J. R. (1969). Conservatism as a predictor of humour preferences. *J. consult. clin. Psychol.* **33**, 271–274.

Manuscript received 28 April 1970

Br. J. soc. clin. Psychol. (1971), **10**, p. 87
Printed in Great Britain

Conformity in Ghana

BY IAN J. McKISSACK*
University College, Cape Coast, Ghana

First-hand contact with African cultural patterns suggests strongly that certain behavioural measures of social processes should show important cross-cultural variations. Specifically, the emphasis on the requirements of social obligation and strong (verbal) support of conventional morality and customs suggest that African subjects would be particularly susceptible to the pressures produced in conformity experiments. Whittaker & Meade (1967) report the results of an Asch-type conformity procedure (one standard and three comparison lines), using college students in the U.S.A., Brazil, Lebanon, Hong Kong and Rhodesia (Bantus). All groups, except the Rhodesians, show 30–35 per cent pro-majority errors, which is the typical finding for this procedure. The Rhodesian group showed 50·7 per cent, yielding a statistically significant difference. The authors qualify their findings because the subjects were volunteers. This qualification is not required in the present author's recent use of the Asch procedure with 26 Ghanaian college students. For each subject there were 18 trials, 12 of them critical trials; a control group made virtually no errors. Five peer-group stooges provided the bogus majority on critical trials. The pro-majority responses were 47·4 per cent of the total critical responses, which is similar to the finding for Rhodesian students. Asch (1956) gives sufficient data for a group of 50 American students, who were tested on the same set of stimuli, to enable a χ^2 comparison of yielding/non-yielding trials for the American and Ghanaian groups, resulting in $\chi^2 = 18{\cdot}1$ ($P < 0{\cdot}001$).

Some discrepancy is introduced by an experiment of Claeys (1967), who used a Crutchfield-type procedure with Congolese students. Again, the judgements were line comparisons. Claeys found no difference between the percentage conformity of his group and two American samples tested by a similar procedure. Two factors are of possible importance. Claeys used 34 neutral trials, and 15 critical trials, a reverse of the usual proportion, which may have reduced the conformity pressure. Also, it is possible, particularly with unsophisticated subjects, that the control panels showing the bogus feedback may not be attended to, or incorrectly interpreted, in the Crutchfield procedure. Claeys reports no check on this point. However, the possibility remains that the type of social pressure to which African subjects are particularly susceptible is that exerted in face-to-face interaction, for which the Asch paradigm is a closer experimental parallel.

Messrs T. J. Powell, T. Boyd and J. K. Nimmo assisted in stage management and post-experimental therapy.

REFERENCES

ASCH, S. E. (1956). Studies of independence and submission to group pressure. I. A minority of one against a unanimous majority. *Psychol. Monogr.* **70**, no. 9. (Whole no. 416.)

CLAEYS, W. (1967). Conforming behaviour and personality variables in Congolese students. *Int. J. Psychol.* **2**, 13–23.

WHITTAKER, J. O. & MEADE, R. D. (1967). Social pressure in the modification and distortion of judgment. *Int. J. Psychol.* **2**, 109–113.

Manuscript received 19 May 1970

---

* Present address: Psychology Department, University of Waikato, Hamilton, New Zealand.

Br. J. soc. clin. Psychol. (1971), **10**, pp. 88–89
Printed in Great Britain

Decision Processes and Obsessional Behaviour

BY A. D. MILNER,* H. R. BEECH AND V. J. WALKER
Institute of Psychiatry, University of London

Obsessional disorder has several features which are impossible to account for solely on the basis of learning theory. The simple avoidance behaviour of phobics is relatively amenable to such explanation, but the complex pattern of doubt, rumination, ritualistic behaviour and repeated checking, which tend to characterize the obsessional patient, is not. One possible common denominator of these symptoms is an abnormality of decision-making, doubt and indecision being characteristic of the process of reaching a decision the consequences of which are of great importance. Ritualistic and checking behaviour might be regarded as the result of an overpowering need for certainty in decisions to terminate quite ordinary activities, while ruminations could be held to represent an 'internalized' ritual. If such a view has any validity, it should be possible to demonstrate and explore, in simple laboratory situations (see, e.g. Audley, 1964), the presence of some decision-making abnormality among obsessional patients.

In the present exploratory investigation, a group of six patients with obsessional symptoms, and a control group of eight patients without such symptoms, were tested. All subjects suffered from depression, and the groups were matched for age. Subjects were required to detect faint auditory tone signals in white noise, under two conditions. In condition A (50 presentations) a 'yes' or 'no' response was required on all trials; in condition B (100 presentations) trials were repeated on request until the subject was ready to respond 'yes' or 'no'. Tones were delivered on 50 per cent of trials on a random schedule.

Using the methods described by Green & Swets (1966), independent measures of signal detectability (d') and response bias (β) were computed from the data of condition A. On neither measure did the groups differ. However, response bias and age were correlated significantly in both groups. The value over all subjects was $r = +0{\cdot}786$ (d.f. $= 13$; $P < 0{\cdot}001$), indicating an increasing reluctance to report signal occurrence as age increases (see also Craik, 1966). In condition B it was found, as expected, that the number of requests for 'repeats' of trials was much higher in the obsessional group (average 0·277 per trial) than in the control group (average 0·065 per trial). The difference was significant ($t = 2{\cdot}35$; d.f. $= 12$; $P < 0{\cdot}05$).

The level of accurate detection under condition B slightly favoured the obsessional group (83·0 per cent correct as against 76·7 per cent), but not significantly so.

The greater tendency for patients with obsessional disturbance to request repeat observations cannot be attributed to lower detectability, as indicated by the results of condition A. Following the signal-detection theory model of Swets & Green (1961) and Swets & Birdsall (1967), the difference might be the outcome of a lower subjective 'cost' of further observations relative to the value of correct responses and cost of errors. It is hypothesized that in obsessional disorder decisions may be deferred (in favour of gathering further relevant information) to an abnormal extent (see also Walker & Beech, 1969).

REFERENCES

AUDLEY, R. J. (1964). Decision-making. *Br. med. Bull.* **20**, 27–31.

CRAIK, F. I. M. (1966). The effects of ageing on the detection of faint auditory signals. *Proc. 7th Int. Congr. of Gerontology, Vienna*, pp. 145–147.

GREEN, D. M. & SWETS, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

SWETS, J. A. & BIRDSALL, T. G. (1967). Deferred decision in human signal detection: a preliminary experiment. *Percept. Psychophys.* **2**, 15–28.

---

* Present address: Psychological Laboratory, University of St Andrews, St Andrews, Fife, Scotland.

Swets, J. A. & Green, D. M. (1961). Sequential observations by human observers of signals in noise. In J. A. Swets (ed.), *Signal Detection and Recognition by Human Observers.* New York: Wiley, 1964.

Walker, V. J. & Beech, H. R. (1969). Mood states and the ritualistic behaviour of obsessional patients. *Br. J. Psychiat.* **115**, 1261–1268.

Manuscript received 6 October 1969

Revised manuscript received 27 January 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 90–91
Printed in Great Britain

The Use of the MMPI for Personality Study of Singapore Students

BY Z. N. KADRI

University of Singapore

For the first time the original English version of the Minnesota Multiphasic Personality Inventory (MMPI) was administered to Asian undergraduates in the Republic of Singapore. Although the survey was carried out among young men and women educated in a Western style, 15 out of the 566 questions considered to be highly colloquial were simplified for the sake of maintaining maximum validity. For instance, 'I very seldom have spells of the blues' was changed to 'I very seldom have spells of sadness or depression' and 'I am often sorry because I am so cross and grouchy' to 'I am often sorry because I am so cross and irritable'. Furthermore, as the majority of students were non-Christians professing other Asian faiths, it was considered necessary to alter questions pertaining to religion. For example, 'I go to church every week' was replaced by 'I go to a holy place of worship regularly' and 'I read the Bible several times a week' by 'I read my holy book several times a week'.

The inquiry was undertaken over a period of six weeks during the 1968–69 academic session when the total university undergraduate enrolment was 3358, of which 76 per cent were men and 24 per cent women. The sample was random in that every third subject, from amongst those who visited the Health Centre during the stated period for preventive inoculations or such trivial complaints as common cold and skin infections, was picked to complete the inventory. All were given the booklet forms of the MMPI and answer sheets and the purpose of the inquiry and the procedure was explained. They were told that rapid answering was better than long deliberation. They were also advised to avoid equivocal responses. Of the total number thus approached, 18 did not consent on the ground that they were too busy with academic work or extracurricular activities. All inventories were scored by hand by a specially trained member of the Health Centre staff. The records of 11 students with 20 or more unanswered items were excluded from the study; but a few with deviant scores on validity scales were not eliminated because the purpose was to ascertain the 'real average' profiles.

The final sample of 260 students (200 male, 60 female) was representative of the prevalent ethnic distribution in the university, i.e. 85 per cent Chinese, 10 per cent Indians or Ceylonese and 5 per cent Malays and others. The age-group they belonged to ranged from 19 to 23 years and they represented a cross-section of all the courses at the university.

The data presented in Table 1 in two separate columns for men and women depict the mean *T* scores without *K* corrections. Also included in the table are results of a similar MMPI study carried out on the California College freshmen in 1952 (Clark, 1954). Although the California data did not include mean readings on L, F, K, Si and Es scales, the striking feature emerging from the comparative study was the great deal of similarity in personality profiles of the two groups. On the clinical scales D, Hy, Mf, Pa and Ma, our men scored slightly higher than their counterparts in California; while our female undergraduates recorded higher mean readings on D, Mf, Pt, Sc and Ma scales than the American freshwomen. Composite scores for our female students were lower than those for men on all scales except Pt, Sc and Si scales. The Singapore men's standard deviation was highest on the D scale, whereas that of women was maximum on the Pd scale.

That personality tests such as the MMPI are expected to show large cultural and subcultural variations is well known. Besides, as illustrated by Clark (1954) and Welsh & Dahlstrom (1956), even the norms may vary in different normal populations. Despite differences in test-taking attitudes and in cultural values and beliefs, the mean scores of Singapore students were in many respects identical with those produced by college students in California. In any comparative study of this nature a number of factors, though difficult to identify, can account for their similarities or differences. The most plausible explanation of similarities between the personality profiles of the American groups and ours lies somewhere between the common knowledge of socially desirable answers to the MMPI queries and the process of internalization of affective–cognitive norms of Western culture as postulated by Fong (1965) in the case of the Hong Kong youth.

Table 1. *Mean standard scores on MMPI scales*

		University of Singapore students				California college students (1954)			
		Men ($n = 200$)		Women ($n = 60$)		Men ($n = 707$)		Women ($n = 763$)	
		Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
	L	53·92	8·22	52·96	7·73	—	—	—	—
	F	57·34	8·68	57·22	8·81	—	—	—	—
	K	55·46	8·73	52·07	7·40	—	—	—	—
1	Hs*	52·54	9·18	49·31	7·02	52·54	8·25	49·69	5·91
2	D	59·98	12·41	56·84	9·76	53·05	9·96	49·69	8·01
3	Hy	57·26	8·92	52·47	9·32	56·26	7·68	54·46	7·20
4	Pd*	55·07	12·04	54·04	12·14	55·90	9·99	54·67	8·73
5	Mf	59·54	8·86	53·00	8·80	58·03	11·61	49·54	9·33
6	Pa	54·31	9·66	46·02	8·42	51·22	7·59	52·72	7·41
7	Pt*	53·15	10·37	53·78	10·50	54·55	9·27	53·17	7·41
8	Sc*	55·10	10·75	55·65	10·45	55·45	9·27	53·98	7·11
9	Ma*	57·75	10·29	56·38	11·44	56·59	10·14	54·07	9·72
0	Si	53·15	8·70	59·78	9·69	—	—	—	—
	Es	52·58	9·81	50·89	8·35	—	—	—	—

* Without *K* corrections.

The author wishes to express his appreciation to the Statistics Unit of the Social Medicine and Public Health Department of the University of Singapore for assistance in processing the data.

REFERENCES

Clark, J. H. (1954). The interpretation of the MMPI profiles of college students: mean scores for male and female groups. *J. soc. Psychol.* **40**, 319–321.

Fong, S. L. M. (1965). Cultural influences in the perception of people: the case of Chinese in America. *Br. J. soc. clin. Psychol.* **4**, 110–113.

Welsh, G. S. & Dahlstrom, W. G. (eds.) (1956). *Basic Readings on the MMPI in Psychology and Medicine.* Minneapolis: University of Minnesota Press.

Manuscript received 13 *November* 1969

Revised manuscript received 9 *March* 1970

Book Reviews

Automation and Behaviour: A Social Psychological Study. By J. K. Chadwick-Jones. London: Wiley-Interscience, 1969. Pp. xi+168. 42*s*.

It is sad to begin a review thus, but this author has been ill-served by his publishers. This is the most difficult book to read which has been encountered for a long time, not because of its content but because of its format. Within an attractive dust-jacket and hard cover are 168 pages of poor quality paper set in a type which seems designed to produce eye-strain and visual discomfort. And lest this be thought a quirk of the reviewer, the book has been shown to others for their spontaneous comments and their reactions have been similar. It seems to be a matter of uneven spacing between the letters of many words in the main text, and unpleasantly small type for the inserted quotations from interviews and the notes on the references at the end of each chapter, which makes it so hard on the eye. But what would be excusable in a war-time production is not good enough for a two-guinea book published in 1969 and one fears it will put off many readers.

Dr Chadwick-Jones deserves better because his subject is an important one. There are few well-documented studies of the effects of automation on jobs, attitudes and behaviour in British industry and so this latest addition to our rather scant literature in this area is to be welcomed. It describes the consequences of technological change in the tinplate section of the steel industry whereby hand operations were superseded by a continuous-flow process requiring different skills and new job roles. The pre-change situation was studied in 1956–57, when a sample 47 workers engaged in rolling and tinning operations were interviewed. After the introduction of the automated system 51 employees who had been transferred from the old factory to a new plant were interviewed. This was at the end of 1962, with some follow-up interviewing in 1963, the men concerned having moved to the new work situation some five or six years previously. In addition, observations were made of work performance and interpersonal and group behaviour over a considerable period at each stage of the investigation, including supplementary studies which continued until 1966. It is therefore a report of extensive research carried out over a number of years.

After an introductory chapter outlining the nature of the investigation, the pre-automated technology is described, and this is followed by three chapters on the pre-change situation in terms of the characteristics of work groups, community ties, job satisfaction and anticipated outcomes of the changeover. This leads to a chapter on the new technology and five more dealing with the attitudes and behaviour of individuals and groups after the change. Of particular interest is Chapter 8 on 'The Enduring Structure of Dissatisfactions', showing that although some initial difficulties and dislikes were overcome as learning took place, others tended to persist, including the disruption to established weekend activities caused by a continuous, rapidly rotating shift system. The men's wives, however, found the new shift arrangements an advantage in that their husbands were more frequently at home and sharing in the life of the family than on the former weekly shift pattern.

The before-and-after situation of supervisors and managers forms the subject of Chapters 12 and 13. These are based on the observations of the research worker and the perceptions of these roles shown in the interview data. In this particular technological change there was no evidence of a closer relationship developing between supervisors and operatives, as seen in some other new technologies, but rather the reverse, which leads the author to end the chapter on supervision, not with his customary 'Summary', but with 'Some Questions' on the implications of his findings.

Chapter 14, 'Conclusions', is rather disappointing. It summarizes again much that has been said before and suggests, tentatively and speculatively, how social behaviour might be assessed and planned for in organizational innovation. We could have done with more discussion of the latter and of Dr Chadwick-Jones' findings in relation to other empirical studies. The references are there throughout the book (more from sociology than psychology), but are mentioned simply in passing and an opportunity has been lost in this last chapter to discuss some of the issues in greater breadth and depth. As it stands, it is too condensed for readers unacquainted with the work cited to be able to envisage what could be the appropriate next step. Nevertheless, all those directly or indirectly concerned with the effects of technology on behaviour should read this book.

SYLVIA SHIMMIN

Attachment and Loss. Volume 1: *Attachment*. By John Bowlby. London: Hogarth Press, 1969. Pp. xx + 428. 63*s*.

Attachment is the first volume of a two-volume series, of which the second is in preparation. The volumes cover the same ground as the series of papers published by Dr Bowlby between 1958 and 1963 on the effects of the loss of mother-figure in infancy or early childhood, but there is more extensive discussion of the bases of the theory of instinctive behaviour which he derives from ethology and control systems theory. Volume 1 is concerned with the problems presented in the first paper of the series, 'The Nature of the Child's Tie to his Mother' (1958).

The book is divided into four parts, which are followed by a 17-page review of psychoanalytical literature on the child's tie to his mother, and 28 pages of references. The principal data of Dr Bowlby's research are detailed records of how young children respond to the experiences of being separated from and later reunited with mother. Very little direct use of this material is made, however, and one is left without information on the numbers of children studied, the circumstances surrounding their separations, the nature and duration of follow-up work and the role of the male parent.

The first part of the book sets out the assumptions from which the author starts and compares these with Freud's. The underlying theme is that the views are compatible, the main difference being the abandonment by Dr Bowlby of the retrospective approach usual in psychoanalysis in favour of direct observation of children. Chapter 2 presents a precis of the empirical observations, from those of Dorothy Burlingham and Anna Freud (1942, 1944) to those of Heinicke and Westheimer (1966) and Robertson and Robertson (1967), for which explanations are sought.

The second part of the book is concerned with instinctive behaviour, and an analogy with control systems in engineering is drawn. Ethological material is incorporated in this part; the work of Konrad Lorenz and Niko Tinbergen is cited and the relevance of the concept of imprinting urged.

In the third part, entitled 'Attachment Behaviour', Dr Bowlby deals with the nature and function of attachment behaviour in subhuman primates and in man, and then outlines a control systems approach to this phenomenon.

Part 4 deals with the ontogeny of human attachment. In this section the author denies expressing the view that mothering should always be provided by a child's natural mother, or that mothering cannot safely be distributed among several figures. It seems likely that all who have worked with 'deprived' children will corroborate his findings that the more insecure a child's attachment to his principal figure the more inhibited he is likely to be in developing attachments to other figures. Various studies of the age at which children in different circumstances focus on one figure are cited, and the importance of 'social stimulation' emphasized.

In a concluding section on 'The Regulation of Mothering' Dr Bowlby presents the question, 'How much mothering is good for a child?' and concludes that he should have as much as he seems to want. One wonders if it is heretical to raise the complementary question, 'How much mothering is good for mother?'

Attachment, Volume 1, clearly written and highly readable, is basically theoretical; it is to be hoped that *Loss*, Volume 2, will present detailed empirical studies which will adequately back the theory.

JOY WELLS

Studies in Animal and Human Behaviour, Volume 1. By Konrad Lorenz. London: Methuen, 1970. Pp. 403. 80*s*.

This book is titled somewhat misleadingly as *Studies in Animal and Human Behaviour*. No studies of human behaviour are included; though, since the book is designated 'volume one', the author presumably intends to remedy the lack by reprinting other work in a further volume or volumes. Unfortunately no such overall plan is revealed.

Six papers written by Lorenz between 1931 and 1942 are reproduced in volume one. Of these the well-known 'Der Kumpan in der Umwelt des Vogels' has already appeared in English as 'The Companion in the Bird's World', and forms the corner-stone for much recent experimental work on imprinting. All six papers were translated especially for the book, and five appear in English for the first time. The new 'Kumpan' reads slightly less well than the older version, but this is a carping criticism since Robert Martin, the translator, has managed a good English style. Of the five newly translated papers one is the original account of Lorenz's observations on jackdaws and other corvids, the main gist of which appears in his book, *King Solomon's Ring*. Three papers are concerned either with naturalistic observations of birds or are philosophical deliberations on the concept of instinct. The most important paper to be included is Lorenz's analysis of egg-rolling in the greylag goose. Ideas generated by this analysis have been influential in subsequent animal behaviour study. At the end of the book Lorenz has appended a series of notes, relating to the text of the papers, in which he attempts to clarify points or show how his present views relate to his earlier writings.

Lorenz's stature and contribution to animal behaviour study are themselves sufficient to justify the publication of a collection of his work. English-speaking students interested in the development of animal behaviour study will be grateful for this opportunity to consult vintage Lorenz painlessly. Lorenz, however, feels constrained to offer his own justification for reprinting the work and this he does in the form of an introduction.

On reading the introduction, one is likely to experience the same kind of distress and embarrassment as on witnessing the 'comeback' of a retired boxing champion. The introduction is petulant and ranting. His reason for presenting the essays seems to be to justify views he has expressed, for example, in his recent book, *On Aggression*. These views have been attacked and his intention now seems to be somehow to provide the rigorous scholarship, undoubtedly lacking in that publication, in order to confound his critics.

It may be an exaggeration to say that the disservice to ethology done by *On Aggression* outweighed the service done by all his previous work, but the book certainly did not add to Lorenz's scientific status. The introduction to this book will do nothing for his scientific reputation either. Lorenz thrashes about indiscriminately, knocking down straw men. So keen is he on the idea—sorry, *fact*—that human behaviour is determined to a large extent by instinct, that he would describe any disagreement with this contention as itself an expression of instinct. This is presumably the disagreement instinct. He poses as some modern-day Galileo, assailed on all sides by critics and dissenters, and expresses 'justified satisfaction that it is against myself that the brunt of doctrinaire attack is directed'. His answer to the critics seems to be that they should sit and watch birds and read the works of Whitman and Heinroth.

A fruitful period for animal behaviour study occurred during the 1950s with a coming together, literally in a series of confrontations, of ethology and psychology. It might be said that both sides retired to lick their wounds, each recognizing the other as a worthy contributor, but also realizing that each was generalizing beyond its facts. Throughout the 1960s both sides were busy gathering facts and learning from each other. We have still a great deal to learn. It may be pertinent to point out that after more than ten years of fairly intensive study we are still unable to decide whether imprinting in birds represents a special (instinctive) process or whether it can be accounted for in terms of associative learning. How much less then are we able to apply or deny instinctive labels for particular behaviour patterns of humans.

Some may well share Lorenz's view that our present prospects as a species look decidedly gloomy. It is possible that, by the time psychology or ethology is able to contribute to decisions concerning the desirability or outcome of particular alternative courses of action, it will be too late to prevent blowing ourselves up or irreversibly destroying our environment. But we should not pretend, in the name of ethological or psychological science, that we have all the answers now. If we are to make decisions, it should be clear that they are political ones even if they are made by scientists. What is required from scientists as such is less arrogance and more facts. Professor Lorenz has, of course, contributed handsomely not only to the store of facts but also in ideas which have proved stimulating to others. Some of the best of these are to be found in this book. For anyone reading it there is no reassurance that Lorenz shares recognition of a need for more.

JOHN WILCOCK

Science, Technology and American Foreign Policy. By Eugene B. Skolnikoff. Cambridge, Mass.: M.I.T. Press, 1969.

When it first appeared over two years ago, Eugene Skolnikoff's book was one of the few to deal, in detail, with some of the issues of political science that fall into the area of science and public policy. The recent publication of the paperback edition holds out the opportunity for a reassessment of this volume in light of more recent thinking both on the interaction of the natural sciences with politics and more specifically on the policy process itself.

Skolnikoff's interests focus on the impact of the sciences (essentially the natural sciences) and their associated technologies on foreign affairs. Furthermore, he speaks not only as an academic observer but also as one who served for five years on the White House Staff of the President's Scientific Advisory Committee in the U.S. The objectives of his study are two-fold: 'first to provide that systemic analysis of the relationship between science and technology to foreign affairs that is a necessary prerequisite to understanding; and second to explain and evaluate the organization within selected government agencies designed to reflect that relationship within the policy process'. The emphasis of this book is therefore firmly on U.S. administrative structures and hence much of the latter parts of the text is devoted to detailed descriptions of how the organizational forms adopted by the Executive and State Departments have adjusted (or failed to adjust) to accommodate the increasingly sophisticated demands being made on them by aspects of science and technology.

To the reader only passingly concerned with the intricacies of American governmental structures, but interested in the effect of science on the broader issues of policy, Parts II and III of the volume will hold out more than the later sections. These two sections deal respectively with foreign policy issues arising immediately from the scientific enterprise (e.g. collaboration in space and atomic energy, the necessity of developing international science policies, the problems of possible environmental alteration) or more indirectly foreign policy problems that contain important scientific and technological elements (e.g. arms control, foreign economic assistance, bilateral and multilateral relations).

The author is at his best in underlining how, in many cases, the narrow and isolated viewpoints of the scientists on the one hand and the policy-maker on the other, have resulted either in conflict between the parties or in the non-recognition of the 'total' or systematic nature of problems. Thus, while scientists have sometimes failed to appreciate the political nature of international collaboration (p. 217), too great an emphasis on international programmes by governments may seriously weaken national capabilities (p. 79). A further example comes from the area of development aid—a field where initially the scientific community are observed to have 'underrated the complexities of development' and to have lacked 'any clear conception of the actual role of science and technology in the development process'. The many difficulties involved with this area are well illustrated in the section on technology transfer (pp. 188–95).

However, while ably demonstrating situational complexities, Skolnikoff frequently slips from description into prescription without first establishing the 'systemic analysis' that he sets off to provide. Too often phrases such as 'if . . . then' or the word 'must' introduce solutions which are not so much at error in themselves but rather that through neglect of situational constraints or consideration of alternative modes of behaviour are offered as 'obvious' programmes for action. In contrast to such simplification, however, it should be noted that the debate over the 'technology gap' has involved more complex issues than are ascribed to it here (pp. 126–32), while some of the organizational structures easily accepted by the author as functional have, with the passage of time, proved far from satisfactory (e.g. U.S. co-ordinating mechanisms for their oceanographic programme, p. 246).

One of Skolnikoff's further errors is that his analysis is structured so that the content of scientific activity frequently appears unimportant, yet paradoxically he criticizes an overseas assistance programme for 'lack of sufficient understanding of science itself and of how the scientific community operates' (p. 294). Yet even more puzzling is that, given the experience of the author in science policy at the executive level, the text is largely devoid of personal insights and of information other than that readily obtainable from secondary sources. An exception to this criticism is the discussion of American involvement with United Nations programmes (see ch. 8); however, frequently there is too little emphasis on policy as a process, and little evidence of any consideration, in depth, of social and psychological factors governing decision-making in these areas. To Skolnikoff rational policy-making with clearly defined goals is seen as 'fortunately' absent (p. 13) and it is perhaps with regard to this area that the book is already somewhat dated. The last two years have seen a vast increase in the literature on policy-making in general and policy analysis in particular. The theoretical continuum of this literature ranges from the incrementalist doctrines of Lindblom to the more future-orientated ideas of Etzioni and Dror; its emphasis is essentially on a systemic approach to complex problems. It is clear that such methods are not without their difficulties, yet they seem to offer greater hopes of achieving an analytical base towards an understanding of policy problems than the somewhat *ad hoc* methods employed in this text.

However, such criticisms should not detract from the qualities of a book, which, in a clear and uncomplicated style, delineates many of the complex interactions of science and technology with the foreign policy process. It is informative and provides novel insights to many of the problems involved. That it sometimes, enticingly, holds out rather simplistic answers to these issues should be taken as a challenge rather than as a mark of failure.

W. I. JENKINS

Personality: A Behavioral Analysis. By Robert W. Lundin. London: Collier-Macmillan, 1969. Pp. ix + 464. 75*s*.

This 'revised and updated' version of *Personality: An Experimental Approach* (1961) boasts approximately 100 new references, most of which deal with recent work on operant conditioning of human subjects.

But while there are many new names on view, Professor Lundin's game remains essentially the same: to define 'personality' out of existence by describing it as the sum total of an individual's learned responses. In a lively and forceful style he presents the gospel according to Skinner, discussing many aspects of the experimental analysis of behaviour in terms of learning theory, and landing out with frequent side-swipes at other points of view.

His enthusiasm is buoyant and should carry along many grateful students who will recognize a rare treat—a thorough exploration of theories and results in an important area of psychology which also is fun to read.

The book's virtue is its occasional vice, since enthusiasm tends to brim into over-optimism. For example, 'experimental neurosis gives us a good deal of understanding about the precipitating causes of abnormal behaviour in humans' sounds like an article of faith rather than a conclusion based on fact. And Lundin's preference for everyday language leads him sometimes to describe at the expense of explanation: discussing the origin of 'escape behaviour' he writes: 'many individuals, the "lone wolves" who walk by themselves, seldom asking advice or counsel from others, have had in their past too many aversive stimuli in the presence and activities of other people'.

Such a statement, which glosses over the really interesting question of exactly how some social stimuli come to be perceived as aversive by some people, labels but does little to explain.

However, academic nit-picking apart, Lundin makes many positive points. His objective interpretations of Freud's theories on anxiety and defence mechanisms are always interesting and sometimes illuminating. While his attempts to draw an experimental analogue do not always convince, his open-minded concern to achieve some integration between views which, superficially, would seem to be poles apart, reinforces one's impression that this book should be required reading for students.

JEFFREY GARLAND

Non-Freudian Personality Theories. By P. James Geiwitz. Belmont, Calif.: Brooks/Cole, 1969. Pp. xi + 147. 20*s*.

Modest aims—'a brief inexpensive treatment of personality theories . . . a small book about what a number of psychologists say about the normal behaviour of adults'—are engagingly realized in this addition to the useful 'Basic Concepts in Psychology' series edited by Edward L. Walker.

Inevitably, the author has made selections—Gordon Allport and Sheldon among them—which are of more historical significance than contemporary relevance. But there are some enterprising choices—Atkinson's studies of achievement motivation and Festinger's of cognitive dissonance—put forward because the ideas raised have generated a good deal of research, although they do not in themselves have the status of general behaviour theories.

One omission—that of Eysenck—is a little surprising. Geiwitz explains, 'since my emphasis is on approaches rather than particular theories, the factor-analytic theory of Eysenck was not included, whereas Cattell's similar approach was', but this seems lame, since he has found room for both Miller & Dollard and Bandura & Walters, although these could be said to have a 'similar' approach.

Perhaps Geiwitz's ultra-cautious line on factor analysis takes us to his more basic reason for shelving Eysenck and giving Cattell relatively little space. He complains, 'there is no question that such techniques are often applied blindly and inappropriately, with results that are somehow impressive but meaningless', and suggests that our present trait constructs, to which factor analysis has contributed, do not appear to account for much variation in human behaviour. This view, one feels, takes us beyond scientific reserve and over the brink into downright pessimism, minimizing the positive gains that have been piled up by factor-analytic research.

But, in general, this book represents an important job well done. Year after year, students of personality theory have ploughed through Hall & Lindzey and similar tomes, cradling their steaming brows and muttering 'there must be an easier way'. There is, and all credit to Geiwitz for having signposted it so well.

JEFFREY GARLAND

Brain Damage and the Mind. By Moyra Williams. Harmondsworth: Penguin Books, 1970. Pp. 175. 7*s*.

This paperback provides a wide view of the field, lucid definitions of most of the main patterns of impairment attributable to brain damage, and stimulating speculations about their physiological basis.

It has the undoubted advantages of shorthand (for the reader trying to absorb a mass of clinical and experimental data); but the unavoidable limitations also. Memory and perceptual problems are generously reviewed; spatial and topographical disorders receive less attention. The final comparison between the effects of focal lesions and dementia perhaps assumes more homogeneity in the focal lesion group than the evidence warrants, and some of the figures are controversial (cf. labelling on Fig. 11, p. 113), or misleading (cf. printer's error on Fig. 5, p. 70). In general, however, the student of neuropsychology will find this book a helpful introduction to the subject and a useful source of reference.

FREDA NEWCOMBE

Foundations of Abnormal Psychology. Edited by Perry London and David Rosenhan. New York and London: Holt, Rinehart & Winston, 1968. Pp. ix + 644. 94*s*.

Lately there has been a welcome increase in the number of textbooks devoted specifically to 'abnormal psychology' (or 'psychopathology' as some authors prefer to call it). This is not to say one wants to see a proliferation of such books, a common affliction in some areas of psychology. But when, for so long, we have had to make do with textbooks of psychiatry masquerading as psychology, or books with a narrow orientation presuming to represent the whole field, or even collections of papers (albeit valuable) pretending to be textbooks, it is heartening now to be able to refer to and recommend to students a reasonable range of genuine textbooks in abnormal psychology. This is one. The others are *Psychopathology* by A. H. Buss (Wiley) and *Principles of Psychopathology* by Brendan Maher (McGraw-Hill).

Foundations of Abnormal Psychology, in a review elsewhere, has been rather dismissed as an 'overweight lightweight'. One needs to be clear, however, as to the readership of such a book before dismissing it. When welcoming textbooks giving a comprehensive coverage of abnormal psychology this reviewer had in mind (1) the psychology undergraduate seeking information on one area of an over-full curriculum; (2) the postgraduate trainee clinical psychologist needing more detailed treatment of and reference to the field of abnormal behaviour; (3) the trained professional needing to keep abreast of his subject; (4) representatives of related professions aiming to keep informed (or even become experienced themselves, e.g. DPM candidates). Any other group to whom the book might have appeal would be regardable as a bonus. Within such a readership classification London's and Rosenhan's book might be regarded as having the edge over the other two for group 1. Buss's and Maher's books, however, might win in groups 2 and 3, while there would not be much to choose for group 4: each book has advantages *and* disadvantages for this readership. London and Rosenhan, however, would probably collect more bonus points from biologists, sociologists, nurses and the so-called intelligent layman or fringe professional.

This does not make London's and Rosenhan's book a 'lightweight' to this reviewer, who regards it as a commendable asset to have begun at the beginning and tried to set down some basic concepts in simple terms (some admirable visual aid diagrams, however, do tend to be offset by some quite superfluous and naive photographs and 'pop' colour-plates). From such a starting point, however, it has taken the authors 140 more pages than Maher and 180 more than Buss to get to the end, and even then it is questionable whether they have got as far or seen so much on the way. This is partly because London's and Rosenhan's contributors have often become rather wordy and protracted (inevitable perhaps in calling upon a wide authorship but undesirable when catering apparently for a mainly undergraduate or lay audience); and partly because *Foundations of Abnormal Psychology* approaches abnormality from the statistical viewpoint and deals with 'positive abnormalities' such as genius and creativity as well as the more traditional 'negative abnormalities'. At first sight this seems to be a good idea and allowable if 'psychopathology' is excluded from the title. On reflexion, however, it seems to serve little purpose: firstly, the exceptional or useful deviations from the statistical average tend to be dealt with these days by other psychology books, and secondly they are hardly the concern of the bulk of the inferred readership of this book who will be working upon problems of the social casuality (whether prisoner, patient, parishioner, pupil or client).

Thus *Foundations of Abnormal Psychology* is an overweight (and hence overpriced) book for its essentially simple and basic conceptual approach. This is a pity because it contains a great deal of useful material. London and Rosenhan themselves show in their chapters a rare flair for expressing ideas entertainingly as well as lucidly. Hirsch and Hofstetter contribute a useful second chapter on 'Behaviour Genetics'—an area of ignorance or out-of-date information for many psychologists. Several other names are familiar to British readers, such as Ross Ashby, Bandura, Treisman and Brendan Maher himself, who has taken time off from his own series of books to contribute the third chapter on 'Behaviour Dynamics'. Kagan contributes chapter 4 on 'Personality Development' and then comes the section on 'Positive Abnormalities' already discussed. This takes one to the half-way stage of the book, leaving just over 300 pages for the negative abnormalities. It is the conceptual framework of this book that these tend to be covered by an excess of social and learning explanations to the exclusion of classificatory description and other theoretical discussions. Hence the book's likely appeal to those chiefly concerned to get a broad view of the field. Yet a great deal of up-to-date work is included as, for instance, that on arousal. No mention is made of abnormal behaviour consequent upon brain injury, however, nor does this come into Buss's scheme of psychopathology. Even Maher devotes only one very general chapter to this. Perhaps this is part of the current endeavour to establish that criteria of abnormality should not be restricted to medical concepts of 'illness'. Yet abnormal behaviour arising from trauma or infection seems to this reviewer to be just as legitimate a field of psychological inquiry as abnormal behaviour arising from genetic defects or adverse environmental experiences in childhood.

It remains to mention one outstanding feature of excellence which will commend this book to any student concerned to revise in a hurry, or to any researcher eager to find his way readily into the literature. This is the summary to be found at the end of each chapter, together with a glossary of terms used, a list of references, *with* titles (not simply journal sources), and a list of suggested further readings with a thumbnail sketch of what each is about. For this reason, for its readability, and because of the paucity of textbooks in this field, London's and Rosenhan's book should find its way on to the shelves of university and many other libraries. Buss's and Maher's books, however, will probably appear more frequently on personal bookshelves.

A. D. BLACK

Changing Man's Behaviour. By H. R. Beech. Harmondsworth: Penguin Books, 1969. Pp. 282. 8*s*.

Beech presents a fair and balanced discussion and evaluation of one of the newer and exciting developments in the art and science of helping human beings, that of behaviour therapy. The book can be read and understood by a general audience as well as being pertinent and useful to a professional practitioner. The past few years have seen the publication of a plethora of research papers, case studies and books on behaviour therapy. Beech's book holds its own very well with at least the half-dozen other recent behaviour therapy treatises in scope of coverage, comprehensibility and illustrative case studies. Despite all that has been published about behaviour therapy, its definition and parameters are still in dispute even by its advocates. Beech takes a definitional middle road in referring to behaviour therapy as 'a collection of techniques'. He makes quite clear that these techniques, based in large part on learning theory and other experimental psychology results, offer a very real alternative, not an adjunct, to traditional psychotherapy. Beech presents the by now almost standard behaviour therapy techniques involving desensitization, relaxation, aversive procedures, operant conditioning, assertive training and self-control. Modelling techniques and token programmes in institutions and schools are not as well represented in the book.

Clearly Beech is a proponent of behaviour therapy but not a cultist. In fact, throughout Beech maintains the attitude of detached objective observer as he seeks to evaluate the evidence. He devotes an entire chapter to 'the case against behaviour therapy' primarily in response to the Breger and McGaugh critique. Throughout Beech's evaluation hinges on research findings, as it should. Many of the theoretical weaknesses of behaviour therapy and the as yet unclear linkage between the laboratory and clinical application are clearly explored by Beech. The view presented is that of a behaviour therapist who is basically an applied scientist emphasizing 'a critical approach to his own theories and treatments, so that he is constantly asking questions about his failure to provide help rather than seeking evidence which confirms an existing viewpoint'.

Beech, however, does not give a complete view of current behaviour therapy. Others may emphasize additional factors such as its linkage to more general behaviour influence procedures derived from social psychology, the implications of a social behavioural model of deviant behaviour as opposed to the more traditional disease model of psychopathology, and the thrust of behaviour therapy beyond the one-to-one relationship into the community and into social institutions.

One can only share Beech's frustration in his conclusion that research in this area of behaviour therapy is still 'piecemeal, small-scale, and inevitably narrow' in its scope because of low social priority. Fortunately, this well-written and enjoyable book is likely favourably to influence lay and professional attitudes, appearing as it does in the excellent Penguin series where it should reach a wide audience.

LEONARD KRASNER

Psychological Diagnosis in Clinical Practice. By Benjamin Pope and Winfield H. Scott. New York: Oxford University Press, 1967. Pp. x+341. 81*s*.

The title of this book has a continuation—'with applications in medicine, law, education, nursing and social work', which is rather misleading, as these activities are only mentioned here and there, in passing. If members of these professions looked into this book for information as to how psychological tests were specifically related to their problems, they would be disappointed. It is in fact a description of the practice of clinical psychology, associated in the usual way with psychiatry and neurology.

It is an account, clearly, of the authors' own methods, and confines itself to a few assessment techniques, which is probably preferable to books which list the baffling superabundance of tests which are available to clinical psychologists. The first half of the book is called 'Theoretical Foundations'. This consists of two general chapters on the function of clinical psychologists and the sort of problems they are presented with, a chapter on psychometric techniques, mainly the Wechsler intelligence scales and the MMPI, and a chapter on projective techniques, mainly the Rorschach. These two chapters contain exceptionally clear accounts of the tests, and there is a very good section on types of pathological thinking as shown on the Rorschach. The second half of the book, 'Clinical Applications', also consists of four chapters, dealing with evaluation in psychoses, neuroses, etc., 'central nervous system impairment' and forensic psychology. These chapters consider the types of problem which arise in each category, and are well illustrated with case histories and detailed test results. The authors keep their feet well on the ground, and are always concerned to make statements which will be of some practical value to the patient.

The good thing about this book is that the case records are not presented as typical productions of schizophrenic or depressed patients; such records tend to be textbook cases which the psychologist never sees. They are presented as answers to actual questions posed by the referring agents, e.g. 'What are the relative weights of organic and functional factors in this patient's illness?' The authors take us through the process of analysing the question, interview, interpretation of test results, to the eventual report. To those of us who feel that American clinical psychologists have greater scope and executive control of the patient than we do, it is comforting to note that these psychologists, at least, seem to work in a situation very similar to ours, with similar problems. Also, they do not shrink from quoting cases where they were forced into fence-sitting reports of the kind we know too well.

T. G. CROOKES

CHANGE-OVER TO SI UNITS

From January 1972 the journals of the British Psychological Society will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Further information about SI Units is contained in the revised *Suggestions to Authors* pamphlet issued by the Society and obtainable, price 5*s*. (U.S.A. $1.00) post free, from Cambridge University Press.

Authors who refer to physical measurements in their papers should now use SI Units except where it would be absurd to do so. Conversion tables will be published during the change-over period but in *Br. J. soc. clin. Psychol.* these will be restricted to units of length as others rarely occur. When they do, conversion factors will be stated in footnotes. Common units of time (e.g. hour, year) will persist.

1 metre (m) = 39.37 inches
1 metre (m) = 3.28 feet
0.0254m = 1 inch
0.3048m = 1 foot

Science of Behaviour

the leading
publishing venture
in psychology

Announcing six new titles

Now published

An Introduction to Social Psychiatry
Ransom J. Arthur, Commanding Officer of the Navy Medical Neuropsychiatric Research Unit, San Diego, California 35p

Pattern Recognition
D. W. J. Corcoran, Senior Lecturer in Psychology, Open University 50p

Schools and Socialization
A. Morrison, Senior Lecturer in Education, University of Dundee and D: McIntyre, Senior Research Fellow, Department of Education, University of Stirling 40p

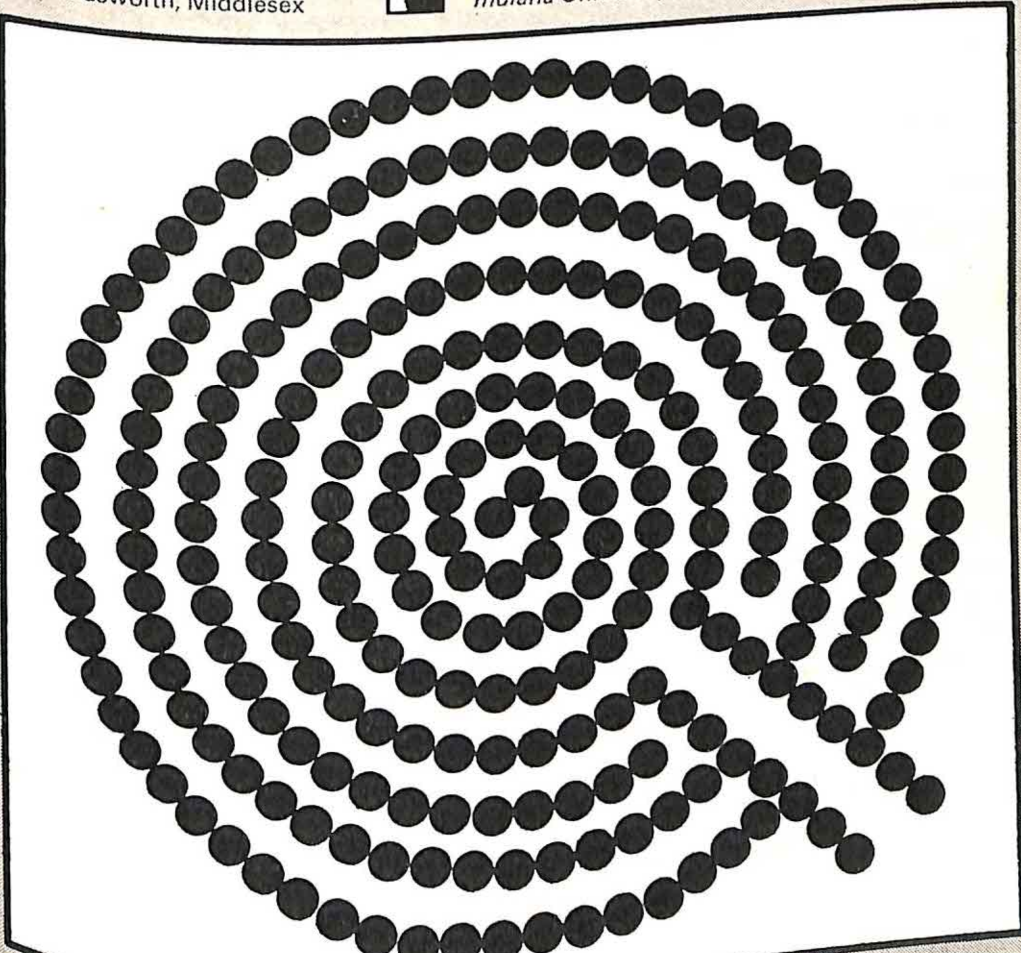
The Study of Twins
Peter Mittler, Director of the Hester Adrian Research Centre for the Study of Learning Processes in the Mentally Handicapped, University of Manchester 45p

To be published in June

Interpersonal Perception
Mark Cook, Research Officer, Department of Experimental Psychology, University of Oxford 45p

Mathematical Models in Psychology
An Introduction
Frank Restle, Professor of Psychology, Indiana University 50p

For further information write to:
Penguin Education,
Harmondsworth, Middlesex



Personality: An International Journal

Vol. 2, No. 1, 1971

CONTENTS

An Investigation of Physical and Personality Factors Involved in Learning to Swim
C. A. Hardy and D. K. B. Nias 1

The Influence of Psychiatric Status and Sex on Semantic Differential Response Style
Peter N. Priest 9

Skin Conductance Responses and Personality Dimensions
Ngaire V. Adcock 15

Context Effects in Response to Affective Stimuli
William Griffitt 23

Personality and Response to Pressure from Peers vs. Adults
Y. Rim and H. Seidenross 35

Cross-Cultural Personality Research and Basic Cultural Dimensions Through Factor Analysis
Robert A. C. Stewart 45

Social Desirability Response Tendency and Self–Ideal Congruence
Bishwa Nath Mukherjee and Ruby Sinha 73

Published quarterly by Plenum Publishing Corporation, 227 West 17th Street, New York, N.Y. 10011 under the auspices of the Society for Personality Research Incorporated.

Subscription price: Volume 2, 1971 (four issues) $12.00.

British Journal of Educational Psychology

Editor: J. D. NISBET

Vol. 41 Part 1 February 1971

SYMPOSIUM ON READING DISABILITY:

1. T. R. Miles, Bangor: More on dyslexia.
2. T. T. S. Ingram, Edinburgh: Specific learning difficulties in childhood: a medical point of view.
3. M. M. Clark, Strathclyde: Severe reading difficulty: a community study.
4. S. Naidoo, London: Specific development dyslexia.

S. B. G. Eysenck, D. K. B. Nias and H. J. Eysenck, London: The interpretation of children's lie scale scores.

P. N. Richards, Bath, and N. Bolton, Durham: Divergent thinking, mathematical ability and type of mathematics teaching in junior school children.

P. Woods and J. R. Hartley, Leeds: Some learning models for arithmetic tasks and their use in computer-based learning.

J. M. Ross, London, and H. R. Simpson, Harpenden: The national survey of health and development.

S. A. Richardson, New York, and A. Green, London: When is black beautiful?

W. M. Bart, Minnesota: The factor structure of formal operations.

A. M. Lucas, South Australia: Multiple marking of a matriculation biology essay.

Research Notes by M. D. Cowell and N. J. Entwistle, Lancaster; P. Kline and A. Gale, Exeter; A. F. Laing, Swansea; A. D. Findlay, Strathclyde; D. McNamara, Lancaster.

Theses Abstracts and Book Reviews

Subscription £4 a year or £1.50 per part to:
SCOTTISH ACADEMIC PRESS LIMITED
25 Perth Street, Edinburgh EH3 5DW

INTERNATIONAL REVIEW OF APPLIED PSYCHOLOGY

The *Review* is the journal of the International Association of Applied Psychology. It is circulated without extra charge to all subscribing members of the Association. Persons wishing to become members of the Association should contact the Secretariat, 47 rue César Franck, 4000-Liège, Belgium.

In 1968 the *Review* was made generally available for sale to non-members of the Association at the cost of £3.00 sterling; 35 French frs.; U.S. $7.00, for the two issues which are published in April and October respectively. Articles covering all branches of applied psychology appear in French or English and each article is followed by a summary in the other language.

Editor: Professor L. S. Hearnshaw, 7 Abercromby Square, Liverpool 7

Vol. 20, No. 1 April 1971

CONTENTS

Le XVIIe Congrès international de Psychologie appliquée

The Long-Range Impact of a Programme of Humanistic Studies for Business Executives on Managerial Attitudes and Behaviour. By Morris S. Viteles

Etat actuel des tests projectifs en France. Par Yvonne Castellan

Psychology in the University of Thessaloniki and elsewhere in Greece. By Roland Harper

Etude psychométrique d'une population d'ingénieurs-techniciens. Par J. B. Dupont, F. Gendre, A. Parisod, et A. Speck

The Concept of Job Satisfaction among African Workers, in particular in a Nigerian Industry. By Mallory Wober

Orders and subscription enquiries to the Periodicals Department
LIVERPOOL UNIVERSITY PRESS
123 GROVE STREET, LIVERPOOL, L7 7AF

Br. J. soc. clin. Psychol. (1971), **10**, pp. 101–113
Printed in Great Britain

Adapting Dawson's Traditional versus Western Attitudes Scale and Presenting Some New Information from Africa

By MALLORY WOBER

Department of Psychology, University of Bristol

This paper extends work on a compound scale of Traditional and Western values, developed by Dawson (1967) in Sierra Leone. Eighteen topics yield scores on westernness of response for each topic; overall subscale scores for approval of Traditional and Western items are also available. Two groups of respondents are manual workers and clerks, in a South Nigerian industry. Differences from Dawson's studies include: some shortening and changes in the scale; 'logicality' in the responses received, in that approval of western items by a group is accompanied by some rejection of traditional ones; lack of whole-scale high correlations between western attitudes and rated work efficiency. It is suggested that the method of presentation differed between the two studies and may have affected results. An element of comprehension may have affected Dawson's studies which could be partly responsible for strong correlations between measures of attitude and work efficiency; this element of comprehension may not have been so salient in the present study. A similarity with Dawson's findings is that some signs of 'resurgence' of less western views is found among some more (western-)educated people. There is also similar evidence of dissimilar attitudes being held at once by certain groups, on certain topics. The overall design of Dawson's device is presented as being very useful and capable of suggesting sensitive pointers to an area of study which may question one of the basic assumptions of balance or consistency theories of attitudes.

Dawson (1967) has developed scales for measuring Traditional, Western and intermediate attitudes among peoples in West Africa. He has related measures of attitude (Dawson, 1963*a*) with work efficiency, with styles of perceptual inference (1967) and with levels of educational attainment and 'exposure to western influence' (1969*a*). In two more recent papers (1969*a*, *b*) Dawson has put forward a 'consistency theory of attitude change'. He suggests (1969*a*, p. 41) that 'T–W attitude change . . . is generally an adaptive process in which the individual tends to reorganize traditional and western cognitive elements to achieve consistency. It is thought that this reorganization process results in the formation of new semi-traditional and semi-western attitude objects, appropriate to the changing social context. It is a basic premise of this theory that attitudinal inconsistency will be maladaptive to the individual and this will result in pressures to reduce or eliminate inconsistency.'

Thus Dawson commits himself to a form of cognitive balance theory of attitudes; and he locates this theory within the tradition implied in the works of Heider (1944) and Festinger (1957). He interprets his data as broadly supporting this type of theory, which embodies an assumption that if an individual harbours two attitude-positions on a given topic, which can readily be seen by a logician to be mutually contradictory or discrepant, then the individual will resolve the dilemma. Dawson adds the assumption that as well as this inner (apparent) drive for consistency, there will be a logic of events or situations, which will constrain an individual to review his attitudes and resolve them in some way. Three alternative kinds of resolution (at least) to the West African T–W dilemma would conventionally be held to be possible. Either a person resolves his attitudes in favour of one distinct attitude position (e.g. Traditional); or he resolves in favour of the other (e.g. Western); thirdly, he may evolve a compromise embodying a new way of reconciling aspects of Traditional and Western attitudes. This third alternative is what Dawson sees manifest in his data.

The present research started as a 'mapping' exercise, designed to describe views among Nigerian workers, along lines suggested by Dawson's work in Sierra Leone. The findings did not, however, closely follow the pattern of Dawson's findings on several points, which will be described below. Further, in examining the present data, a fourth alternative strategy whereby an individual may articulate his attitudes has suggested itself. This strategy, which will also be explained below, may also be discerned to be operating among some of Dawson's West African informants. In some cases, the fourth alternative may provide a better fit in explanation than the interpretations offered by Dawson for his own data, and it also seems appropriate for some of the present data. This raises a question about possible limits to the universality of rationalistic consistency theories.

Dawson's Traditional versus Western Attitude Scales

Dawson has twice described his scales (1967, 1969*a*) which were first developed in Sierra Leone and dealt with eighteen topics. Each topic is one on which it was assumed that a Traditional view would exist (T) different from a Western view (W). (It must be pointed out that what is called a 'topic' here, Dawson called a 'group'; he used the word 'group' to refer to a group of attitude-statements; this usage is not followed here since it is accepted that the term 'group' conventionally applies to people, while *topic* would be more suitable for a set of related attitude-statements.)

Accordingly, sentences were written expressing T and W views for each topic. Further, sentences expressing a Semi-Traditional (Semi-T) and a Semi-Western (Semi-W) view were also produced. Dawson used (1967) an unstated number of African and European residents of Sierra Leone as judges, in what he terms (1969*a*) 'an extreme modification of the Thurstone technique' to scale the items. As he acknowledges, the scaling posed considerable problems; some of these impinge on the conclusions which may be drawn from his findings, and this will be referred to below.

Dawson points out that several scoring indices can be derived from his scales. Each respondent reacts to each of the eighteen sentences propounding traditional views on each topic; his collective score on these items gives his position on the T scale. The same applies for the W scale. Further, the total approval of Western items can have subtracted from it the total approval for Traditional items. This is what Dawson called the 'Range of Variation' score ($RV = W - T$). The RV score, which is here referred to by symbols indicating its composition, namely $W - T$, can be computed as a composite for all eighteen topics; it can also be computed for each topic separately, and topic scores of this type can be seen here in Table 1.

THE PRESENT STUDY

Subjects. Two groups of industrial workers were questioned, all of them being employees of the African Timber and Plywood Company's factory at Sapele, in Southern Nigeria. There were 50 clerks (mean age 29·8, S.D. 6·7) and 173 manual workers (mean age 32·7, S.D. 7·3). Thus the clerks were a significantly younger group than the manual workers ($t=2\cdot66$; $P<0\cdot01$). Education was assessed by a four-point scale reflecting school experience (explained in Wober, 1967*a*), and clerks (3·76, S.D. 0·43) clearly have had more education as a group than manual workers (2·67, S.D. 0·83, using a cut-off-point between 3 and 4, $\chi^2=70$, thus $P<0\cdot01$).

These differences, as well as tested ability and other indices, could all be treated as variables distinguishing the two groups apart. The amount of 'western exposure', which is a variable used by Dawson, was not assessed in simple form here. It might be argued that school experience is one satisfactory way of operationalizing 'western exposure', and in this case the two occupational groups have had very different 'western exposure' with the clerks having had more. However, it could be said that as manual workers have spent more time in the industry (10·1 years) than clerks (7·8 years), in this respect the 'western exposure' was more for the workers.

Nearly 59 per cent of the manual workers were Ibos (distinguished as Eastern and Western Ibos at that time); 30 per cent were of the Edo group of peoples, of the then Mid-West State. Exactly 32 per cent of the clerks were Ibos, while 44 per cent were Edos, and the remainder from a variety of tribes excepting, however, any Northerners. An analysis of attitudes by tribe will not be given here, for two reasons. Firstly, there were no definite and systematic differences on attitudes between tribes; further, these differences could not be ascribed to tribe alone, as tribal groups also differ on education and on a number of other salient variables whose effects would have to be partialled out in a lengthy analysis. Such an analysis can be short-circuited by examining some of the effects of these salient variables directly, which will be done. Secondly, it has been decided not to focus on an intertribal analysis, whose scientific usefulness is not held to justify the possible social cost of displaying unwelcome comparisons in the present political context.

Testing procedures with the subjects were carefully managed and observed, and have already been described at length (Wober, 1967*b*).

The Questionnaire

Originally it was intended to apply Dawson's instrument as closely as possible, introducing only small changes of wording if necessary. It soon became clear that several topics were not relevant in South Nigeria, and that the four item per topic scale (T, Semi-T, Semi-W and W items) was difficult to administer convincingly. It was too long, and people felt that T items and the others for any one topic did not refer exactly to the same topic they were meant to. This objection is not dealt with by a scaling procedure. A host of judges could agree that the T item was truly traditional, the W item western, and the others truly in between; but this could be satisfied with these items referring to different topic areas; or the items could marginally diverge from treating of the same topic. Thus in the example given by Dawson (1969*a*, Table 1) the T item on the topic of brideprice includes the elements 'make marriage secure' and 'bring families together'; the W item for the same topic includes the element 'fit in with modern education and progress'. Thus it is not certain whether responses on this topic really concern only the issue of brideprice, or whether they reflect opinions on other matters as well.

This difficulty in item-writing was unavoidable, and occurred with the items devised for Nigerian use as well. The alternative of providing terse statements was not the favoured style of communication with some of the less sophisticated respondents, for whom a more repetitive sentence was more suitable. Both the T and the W items for each topic had to be made to appear equally worthy views to hold, and to do this some degree of elaboration was necessary. With four sentence items for each topic, it is likely that more extraneous elements to the specific topic would creep in than if there were three sentence items per topic. In this way it was hoped to make the present items more 'iso-topical' (i.e. referring to one topic alone), and by this simplification to improve comprehension. So a three item per topic structure was adopted, the items being the Traditional, Compromise and Western sentences.

The present data will concern only the responses for the Traditional and Western sentences. These had been recorded from the original protocols while in the field; but unfortunately the protocols were lost in travel from one university location to another, so the Compromise data are not, nor ever will be available. This is a scientific loss, but does not devalue the task of fully analysing the available information.

The topics included in the present questionnaire omitted four used by Dawson, but included four new ones (here labelled N to Q).

A. Concept of the gift
B. Practice of dowry payment
C. Native *v.* modern medicine
D. The existence of witches
E. Traditional *v.* school education
F. Polygamy *v.* monogamy
G. Women's roles: home *v.* work
H. Congruence of age and wisdom
I. Native customs *v.* modern progress
J. Wifely obedience *v.* democratic roles
K. Rich men's obligations
L. Ownership of discovered property
M. Chiefs, or government to administer tax
N. Women's clothing styles
O. Preference for native *v.* western music
P. Trad. *v.* modern voluntary associations
Q. Vernacular *v.* English as mother tongue
R. Trad. *v.* modern work group preferences

The scoring system was similar to that used by Dawson. For each subject, four scores were derived, including the T, W and the W − T overall scale scores. Further, a score was derived for a person's view on each topic, for example as follows:

	Response scale				
Topic C	Strongly agree	Agree	Not sure	Disagree	Strongly disagree
A sick man should go, or be taken to a native doctor, as they know many things that European-trained doctors do not deal with	+2	+1	0	−1	−2
When someone is sick, it may be something that a native doctor should treat, though often it is best to go to a European-trained doctor	+2	+1	0	−1	−2
A sick man should be treated by a European-trained doctor only, as they have the proper medicines, while native doctors do not treat diseases scientifically	+2	+1	0	−1	−2

A man answering as above will have rejected the Traditional position, and strongly approved the Western one; his score would be for this topic: $+2, -(-1)=3$. This score does not make use of the result from the intervening Compromise sentence, but focuses on the difference between Traditional and Western positions. The possible scores found by this method for any one topic would range from -4 to $+4$. To avoid minus signs the whole scale can be shifted up five points, to run from 1 to 9. The man's score above would then be, for this topic, 8. His component towards the overall T score would be 2 (simply derived by transforming the response scale positions to values 5, 4, 3, 2, 1), and towards the overall W score would be 5. The minimum score on each of the T and W scales, through the whole questionnaire of 18 topics, is 18, and the maximum is 90.

Now it must be pointed out that the T and W scales are not unidirectional scales, each starting from some 'real' minimum level at 18 to a maximum agreement with Western ideas at a score of 90. The actual response scales included the terms Agree and Disagree. Thus it is more correct to realize that the true zero of non-committal opinion lies at a score of 54 (i.e. $18+(90-18)\frac{1}{2}$). A score above 54 denotes overall agreement with T, or with W propositions as the case may be. Scores below 54 reflect disagreement. A T score of 18 means 'disagree strongly' with every traditional item in the questionnaire; 90 means 'agree strongly' throughout. Any group, therefore, which has both T and W scores below 54 has rejected both the T and W attitude constellations; this is a logically consistent position and can be reconciled, especially if one strategy implied in a cognitive consistency theory is correct, with agreement on some intermediate formulation (e.g. Semi-T, Semi-W, or Compromise).

If a group has a T score near or above 54, and a W score also above 54, this denotes acceptance of two 'logically' opposed sets of statements. This is a position that is rationally inconsistent; no matter how emphatic an agreement there also exists with intermediate formulations, this does not alter (and certainly does not 'resolve') the situation that in this case orthodox consistency theory does not account for the facts. It will be seen below that Dawson's data show several instances of such cognitive discrepancies (which are not necessarily evidence of dissonance) at a group level.

The sentence items were produced as follows: they were written with the aid of a Nigerian assistant. Trial questionnaires were given to classes of boys in a secondary boarding school, with whom after the responses had been given the items were discussed. This showed that where T items were rejected, W items were approved, which implies among groups that the items were conceptually distinct. Further group testing with boys of similar age in a government technical college (which however had a much lower academic standard) revealed many examples of 'cognitive disparities' in responses. This suggested cautions were needed over group administration. The sentences were further discussed before being finalized, with a colleague doing a parallel study in the (then) Northern Region of Nigeria. It is accepted here that the question of 'scaling', especially in research on cross-cultural change, is not one of finding some absolute baseline, but of adopting a relativistic position. Judges can be found who will say that groups of sentences embody either a Western or a Traditional view; yet other judges may show a measure of disagreement with this. The present T and W items evoked a substantial measure of agreement between Nigerian and European observers as to their T and W status; yet the normative data show that for a few topics among the groups tested, those who support W items on the whole will support a T item in special cases. This implies that, in these special cases, the scaling has not been valid for the group of people in question. This is not to be treated as a flaw in a methodological perspective, but as a discovery in normative terms.

The administrative procedure used in Nigeria was not the same as had been employed in Sierra Leone. Whereas Dawson (1967) found that men with top Primary Class 6 education or above 'could cope with filling in the scale whether individually or in a group type situation', this was not found to be advisable in the present instance. As an example, one man who had replied to one sentence 'I 'gree small' (pidgin for 'agree', as distinct from ''gree plenty' for 'agree strongly'), on hearing the next item to which he also agreed, changed his first response. He explained: 'when I say "I 'gree small" what I mean is "I no 'gree"'.

This and many other examples showed that comprehension might play an important part in influencing responses; for this reason, all questionnaires were individually supervised with the aid of a Nigerian assistant (who had also substantially helped in wording the items in the first place). No attempt was made to represent the response scale visually with boxes or diagrams as had been done in Sierra Leone. This was tried in pilot investigations, but was found to add complexity, probably in that it involved encoding the answer in an unfamiliar medium. All responses were verbal. Importantly, most respondents said after each topic had been dealt with that they felt their views had been accurately assessed.

RESULTS

Table 1 presents the scores of the two groups of subjects for each topic, and the correlation coefficients between the score for each topic and that for the W – T score for the scale as a whole.

A correlation of the ranks in which the scores from each occupational group may be ordered shows rho = 0·99.

Both occupational groups are seen to present overall 'logically' consistent patterns of response. Manual workers have a T scale score of 48·38, which, being below 54, denotes some disagreement with T propositions; at the same time they approve (W score of 71·48) the western items. The clerks show a stronger rejection of traditional items overall than do manual workers ($t=3·26$; $P<0·01$); and they also approve, but to a weaker extent than manual workers ($t=3·38$; $P<0·01$), the W items. The W−T difference overall (Dawson's 'Range of Variation') is effectively equal in both groups ($t=0·02$; $P>0·1$).

The sizes of the coefficients suggest that omission of topics K, M, H and possibly A would have improved the scales. Some topics, for example C, D, I and Q, correlate strongly among both occupational groups with the overall scale scores.

Table 1. *Mean score on each topic for two groups of respondents, and correlation coefficients between scores on each topic and an index (W − T) of the whole scale*

Manual workers ($n=173$)				Clerks ($n=50$)		
Mean topic score	S.D.	r: W−T	Topic	Mean topic score	S.D.	r: W−T
6·69	1·97	0·21	A. Gift	6·82	1·59	0·25
3·66	2·70	0·28	B. Brideprice	4·24	2·77	0·43
7·16	2·37	0·49	C. Medicine	6·50	2·41	0·56
5·47	2·51	0·26	D. Witches	4·94	2·44	0·49
6·73	2·23	0·36	E. Education	6·24	1·03	0·42
6·75	2·69	0·44	F. Polygamy	7·56	1·70	0·36
7·83	1·45	0·35	G. Women's roles	7·66	1·46	0·31
6·73	2·02	0·29	H. Wisdom and age	7·48	1·63	0·25*
6·32	2·06	0·33	I. Custom *v.* progress	5·92	2·28	0·51
5·40	2·51	0·42	J. Wives' obedience	6·06	1·77	0·19*
2·47	1·79	−0·07*	K. Rich men	3·16	1·99	0·19*
7·85	1·54	0·41	L. Discovered property	8·02	1·44	0·45
7·18	2·14	0·27	M. Admin. of taxes	7·32	1·33	0·19*
6·64	2·29	0·32	N. Women's clothes	5·22	2·12	0·43
5·14	2·39	0·45	O. Taste in music	4·66	2·26	0·37
7·15	1·80	0·34	P. Voluntary assocns.	7·72	1·62	0·29
6·32	2·06	0·43	Q. English language	6·20	1·77	0·53
7·65	1·48	0·39	R. Work loyalties	7·60	1·79	0·37
			All topics:			
48·38	9·16	−0·85	T items	44·56	8·27	−0·86
71·46	6·66	0·80	W items	67·98	7·27	0·70
23·08	12·38		W−T	23·42	12·87	

* Values of r *not* reaching significance at the 0·05 level.

Topic K, on how rich men should behave, is shown among both groups to have been scaled wrongly as far as they were concerned. Yet judges had agreed that the sentence intended as T ('a rich man should help . . . by sharing his wealth with the poor in his family') represented a 'Traditional' point of view fairly; while the W sentence had been judged as representing a 'Western' view ('a rich man's money is his own, to spend on himself, wife and children'). Yet this is a case where not only people who favour mostly T items over W ones favour this T item; but also men who usually favour other W items prefer this T one. This accounts for the low mean score in both groups on this topic.

This could be treated as a simple error in scaling. Yet it is felt that this is to prefer the goal of reliability to that of validity. A glance at the previous topic, J, on whether wives should be obedient or democratic, may illustrate this point. For the workers, a good correlation suggests that the item was 'properly' scaled, i.e. it evokes responses in line with those on other topics; but for the clerks there is a low and non-significant correlation with the overall scale score. This suggests that scaling is in order for one group, but not for another; thus it is necessary to take a relativistic view over scaling.

In general, the scores show that both groups agree more with Western than with Traditional formulations, as for workers there are only two topics with scores below 5 (the mid-point of the scale, at which approval for T or W is balanced); for clerks four topics earn scores below 5, which derives from what was shown earlier, that clerks are less emphatic in support of W items.

Both groups favour a western view on medicine. The company has a good clinic, whose value was highly respected. Both groups had a less distinctly western view about witchcraft. Here is a topic on which there is evidence for people to contemplate from the traditional culture; while a western industry has little to say, either positively or negatively, about witches. Thus traditional beliefs may continue unconfronted with some explicit western alternative. Both groups prefer a more traditional view on the question of dowry or brideprice payments, though this is not an emphatic position.

There is some evidence here that attitudinal inconsistency (though it is not manifest overall) may occur with specific topics. A man who accepts a T proposition and also a W one is in some way inconsistent; his score on that topic would be 5. Another whose replies to T and W items differ by only one position on the response scale, is also inconsistent (unless both his replies consist of rejection); this amount of inconsistency will result in topic scores ranging from 4·0 to 6·0. Among manual workers three topics showed scores between these limits, and there were five such instances among the clerks. For these topics, many people are either agreeing equally with T and W items; or they are equally non-committal on both; or they disagree with both. Several people must therefore be showing cognitively inconsistent response patterns. This was in fact quite clear during the interviews, and represents attitudes which people were firm in maintaining, even when the inconsistencies had been discussed. For this reason if for no other, it seems fair to claim that incomprehension is likely to have been only a small element in determining the present results.

Dawson (1963*b*) maintains that men holding more traditional attitudes tend to be less efficient workers. The present results do not support Dawson's pattern of findings.

Quite clearly, work efficiency does not correlate significantly with the W−T measure, nor with the T or the W measures separately. Several reasons might account for this. The work efficiency ratings could have been non-comparable between the two studies, though the accounts of the rating scales (Dawson, 1963*a*; Wober, 1967*a*) suggest that they should be comparable. That the present ratings are not just a random set of marks is shown by the fact that there are correlations between efficiency and a tested ability (for workers, with RFT, $r = 0{\cdot}26$, $P < 0{\cdot}05$; for clerks, with EFT, $r = 0{\cdot}43$, $P < 0{\cdot}01$). Secondly, Table 1 shows that different topics do not contribute uniformly to the total scales, and the same unevenness exists in Dawson's scales, as is finally disclosed in his recent (1969*a*) study. These discrepancies between the scales could account for the different correlations with work efficiency. Thirdly, there may have been an element of miscomprehension affecting Dawson's results, though not the present ones.

The measure which Dawson is correlating with efficiency is his 'Range of Variation' (W − T here). Dawson believes that when this entity is large, this shows a preference for Western views, which is true. But he also interprets an RV score of close to zero (as with the Arunta; Dawson 1969*c*) as 'indicating High Traditional-Low Western responses'. This is not a valid interpretation, as his Fig. 1 in that paper shows. A low RV score means that those people have a sympathy to Western views and an equal sympathy to Traditional ones. From this example, as well as elsewhere in his work, we see that high RV scores denote people who have resolved their views in favour of Western ideas; while low RV scores denote people who are still in states of dilemma about many matters. These inconsistent response patterns could result from real dilemmas, or from poor comprehension in group testing situations. The latter would be consonant with lower levels of education, and would give rise to statistical correlations between efficiency and attitudinal scores.

Table 2. *Correlations between ratings of work efficiency, and T, W and W − T overall scales*

Group	Source of data	n	r: T	r: W	r: W − T
Female nurses	Dawson (1963*a*)	16	—	—	0·65
Labourers	Dawson (1963*a*)	30	—	—	0·49
Artisans	Dawson (1963*a*)	79	—	—	0·60
Apprentices	Dawson (1963*a*)	159	—	—	0·41
Clerks	Dawson (1963*a*)	37	—	—	0·70
Manual workers	Present study	173	0·06	0·05	−0·01
Clerks	Present study	50	0·07	0·06	−0·01

With the present results, great care and individual testing throughout was used to minimize difficulties in comprehension. This would mean that low W − T (RV) scores would reflect true dilemmas of attitude (or perhaps compartmentalized and insulated attitude sets). There are no good reasons to suppose that educational experience should affect efficiency if the employees are placed in jobs suited to their educational level; in fact zero correlations between education and efficiency will reflect a good company personnel function. The correlations found between these variables were for clerks −0·11 and for workers −0·12; neither of these are significant. Similarly, for clerks the correlation between education and W − T score was −0·04, while the value for manual workers was −0·006. So in this investigation education, job efficiency and measures of attitudes on the whole scale seem to be quite independent of each other.

However, efficiency does correlate significantly with attitudes on specific topics.

All the coefficients in Table 3 reach significance at the 0·05 level. The first three are in the direction that belief in a traditional view is linked with lower rated work efficiency. The other five values show that men who preferred the western views tended to be rated as less efficient. The question of validating items initially scaled as 'western' is thus again shown to merit further consideration. For example, at face value it seems right to assume that preference for tribal-style education is traditionalist while that for school is western. Yet the correlation between attitude on this educational topic and actual level of education is -0.17 ($P<0.05$); this suggests some tendency for the less educated men to prefer the idea of school education, while those with more schooling evidently think less of it. Is then the western view that which is designated as such by the social scientist; or is it the preferences of the more schooled man; or of the more efficient worker? No entirely consistent and clear answer is possible.

Table 3. *Relationship of manual workers' attitudes on separate topics, with rated efficiency at work*

Topic	r*
D. Belief in witches	−0·21
F. Polygamy *v.* monogamy	−0·16
K. Rich men's obligations	−0·19
E. Traditional *v.* school education	0·16
G. Women's role: home *v.* work	0·16
L. Ownership of found property	0·17
N. Women's clothing styles	0·22
Q. Vernacular *v.* English as mother tongue	0·21

* At $n = 173$, all these values are significant at 0·05.

A particular topic illustrating the complexity of currents of opinion is topic K on the obligations of rich men. The less selfish (acquisitive) traditional view was associated with greater rated work efficiency; but as Table 1 suggests, topic K was negatively correlated with 14 of the other 17 topics. Thus people who prefer the traditional view on K tend to take the modern view on other topics; and on five of these topics there is a positive relationship with efficiency, showing that people who take the western view are the more efficient.

Correlations are available between attitudes on individual topics as well as for the overall T and W scales, against other variables. Dawson (1967) has linked attitude positions with the development of certain styles of perceptual inference and cognitive abilities. The present study included tests of cognitive style and perceptual abilities (Wober, 1967*b*). For each group all the separate topics, and total scales, yielded 126 correlation coefficients against the tests. From this number, for clerks only six were significant at the 0·05 level, and among the manual workers eleven were significant. None of these significant values occurred with the total scale scores against tests. This is not thought to be sufficient evidence for any hypothesis that attitudes were related to cognitive styles or perceptual abilities.

Among clerks age was not related to score on the overall scale for T items ($r = 0{\cdot}002$) or to the W – T scale ($r = 0{\cdot}16$); but age was related to the overall W scale ($r = 0{\cdot}30$). This shows a tendency for the younger men to have less approval for the Western items. Here we can distinguish between support for a traditionalistic attitude, for which there is no evidence here, and shrinking from support for western attitudes, which is discernible among the younger clerks. This parallels findings to a similar effect by Jahoda (1961) in Ghana.

DISCUSSION

Dawson's questionnaire device for investigating attitudes in a society where much change is in progress has been amply acknowledged here. It provided a most useful framework for this study in Nigeria. However, there are clearly major differences in the patterns of evidence, and in the interpretations offered between the Sierra Leonean and the Nigerian studies.

Dawson has proposed (1969*b*) that in a milieu of cultural change it will be adaptive to reorganize attitudes so as to compromise between traditional and western views and thus achieve consistency. 'Lack of achievement' (1969*b*, p. 7) 'of these in between compromises will be maladaptive and result in pressures to reduce or eliminate inconsistency'. (Note that the use of the phrase 'lack of achievement' implies a viewpoint embedded in the western value system, with the danger that different or novel attitude-management strategies become more difficult to recognize.) External pressures will prompt a compromise, he says, if internal cognitive initiative fails to produce one. This, Dawson says, should apply, except where there is no western contact (1969*a*, hypothesis 1) or where there is considerable western contact (hypothesis 4).

Examples are given (1969*a*) of hysterical dissociation among students who had neglected to perform certain traditional rites. Though at school far away from home, they had to be sent off to perform their rites, after which they returned happy. This Dawson presents as a strategy to resolve conflict: 'the above cases where attitudinal conflict had not been resolved must be considered as maladaptive, whilst considerable pressures arose to reduce inconsistency'.

It is held here, however, that Dawson's cases and statistical evidence could be interpreted differently. It remains to be shown that the hysterical students reorganized their attitudes to prefer compromise formulations. On the contrary, it would seem that their final return to school argues for their strong approval of this western behaviour, while their need to go home for rites suggests they also strongly respect traditional ways.

An alternative interpretation is that sets of attitudes may be established which congregate about sets of roles individuals occupy when living in particular cultural milieus (e.g. the town or the village) with some sort of cognitive boundary actively maintained to protect the individual from possible confusion and cognitive dissonances that he might not be able to resolve once they were brought to confrontation. Not enough data is yet available to examine the hypothesis of this 'fourth strategy' directly, though evidence for it can be seen in both Dawson's and the present studies.

Dawson shows (1969*a*, Fig. 1) that among four different groups of male skilled workers, support is greatest for the W items. One of the groups is of 40 Fulbe, who as a group have more conservative attitudes (1969*a*, p. 40) than other tribes. According to hypotheses 2 and 3 (1969*a*, p. 43), where 'traditional subjects' have either 'some' or 'more marked' amounts of 'western exposure', 'inconsistency between traditional and western attitudes would be resolved' into compromise formulations. In fact the Fulbe, like the three other groups, have 'peak support' of W items. In the Fulbe case, not only do they support W items, but they also have a mild measure of agreement for the T items (shown by a score exceeding 54, in Fig. 1). So the Fulbe data do not support Dawson's two hypotheses, though they may be said to support his next hypothesis, which is that where there is 'considerable western contact' there will be a peak W response. Now no way of operationalizing or measuring 'western contact' is given. If Fulbe are supposed to be experiencing considerable western contact (or exposure—the words are used alternatively by Dawson), then as they also belong to a conservative tribal culture, they might well be expected to experience cognitive dissonance. In fact they have the flattest curve of the four groups, which means that their W score is 12 points more than their T score; but the important point to notice is that both these scores are in the direction of approval, and this leaves their condition as one of dissonance. That is, if they actively contemplate their sets of opposing cognitions.

Though this situation is seen most clearly with the Fulbe group, it can also be said to apply for the Mende and Temne. It would seem that they all show evidence of holding cognitively dissimilar views; though this is not the same as saying that cognitive dissonance occurs. Thus it can be argued that many of Dawson's data (1969*a*, Fig. 1; 1969*b*, Fig. 3) do not as he contends support his consistency theory of attitude change; they raise and leave open the question of how subjects manage their cognitive worlds.

Dawson states (1969*a*, p. 43) that a respondent answering 'agree strongly' to all four items of a given topic has presented 'an invalid response'. He says (1967) that such responses were made by illiterate subjects and that where a series of these were made the results were discarded. It is not surprising therefore that his prior formulations should imply (1967) and state (1969*b*) that cognitive dissonance has been resolved by subjects by the time their attitudes are assessed. On the contrary, it raises the question of whether good comprehension was ensured during administration, which had to be individually done with illiterate subjects. And if miscomprehension occurred with individual administration, it is likely also to have happened with group administration. In fact, he later (1969*a*) shows that on certain topics there is manifest a substantial level of 'unresolved conflict'.

The present results certainly show that for entire groups of respondents and for the overall scales, logical consistency of results is observed. However, the focus of interest lies on the conditions where inconsistencies continue to exist. In this aspect of interpretation, as well as in the pattern of results, this study differs from Dawson's. Whereas there is no relation here between W – T scores and occupational group (which implies an educational difference), Dawson reports (1967, Table 9) that W – T scores do relate to educational level. This may well affect the comprehension which subjects with modest educational attainments will bring to group administered questionnaires. It is also held possible that if Dawson's W – T score was measuring in part an element of comprehension, this could explain the high correlations he reported with work efficiency, which failed to be replicated here. While Dawson also related attitude measures with perceptual skills (1967) there was no supporting evidence here.

The results here show that clerks and manual workers have attitudes which balance to a similar extent; the more educated clerks show signs of a lesser enthusiasm for clear-cut western ideas. Attitudes on certain topics relate, though not always in the same direction, with rated efficiency at work, though the net correlations between the whole scales and efficiency are not significant.

CONCLUSION

There has been a school of theories in social psychology, exemplified in the work of Heider and of Festinger, that assume that rational analysis and economic-style accounting will make for efforts to resolve inconsistencies in attitudes. Dawson accepts this theory and in several papers drives to interpret his data accordingly. Recently (1969*a*), by discarding the most cognitively discrepant responses as invalid, and by setting a high threshold criterion of consistency, Dawson argues that there is 'only' 24 per cent of unresolved conflict expressed in responses to the Traditional–Western attitude scales he devised.

People in West Africa who work in schools or in industry, but who on the other hand retain strong roots in tribal cultures are certainly exposed to problems of how to manage attitudes which their two contexts may each strongly evoke, but which may conflict. Resolution of this dilemma may certainly occur in terms of accepting Western over Traditional views; or vice versa. Resolution may also occur by fastening on compromise formulations at the expense of clear-cut polar attitudes.

However, Dawson's and the present data also suggest that a fourth strategy of attitude management may exist. This is implied in the evidence that some people agree with Traditional and with Western attitudes, apparently in conflict. If they also affirm compromise formulations this does not remove the fact that they harbour attitudes that rationally are not consistent.

Two questions are therefore raised. The first is as to how cognitively disparate, but enduring attitudes may be successfully managed. To answer this it may be relevant to examine role-setting and how this might enable separate sets of attitudes to be compartmentalized and insulated from each other. The second question is that if this strategy occurs in West Africa, it probably also occurs in Western countries and studies devoted to investigating such a phenomenon would be enlightening. This is one of the important functions of cross-cultural research, to learn more about ourselves by studying other peoples. We may either find that a phenomenon exists among us; or we may be more certain that it does not.

REFERENCES

DAWSON, J. L. M. (1963*a*). Psychological effects of social change in a West African community. (Unpublished D.Phil. thesis, University of Oxford.)

DAWSON, J. L. M. (1963*b*). Traditional values and work efficiency in a West African mine labour force. *Occup. Psychol.* **3**, 209–218.

DAWSON, J. L. M. (1967). Traditional versus Western attitudes in West Africa: the construction, validation, and application of a measuring device. *Br. J. soc. clin. Psychol.* **6**, 81–96.

DAWSON, J. L. M. (1969*a*). Attitudinal consistency and conflict in West Africa. *Int. J. Psychol.* **4**, 39–54.

DAWSON, J. L. M. (1969*b*). Theoretical and research bases of bio-social psychology. *Univ. Hong Kong Gaz.* **3** (suppl.), 1–10.

DAWSON, J. L. M. (1969*c*). Exchange theory and comparison level changes among Australian Aborigines. *Br. J. soc. clin. Psychol.* **8**, 133–140.

FESTINGER, L. (1957). *A Theory of Cognitive Dissonance.* New York: Harper & Row.

HEIDER, F. (1944). Social perception and phenomenal causality. *Psychol. Rev.* **51**, 358–374.

WOBER, M. (1966). Psychological factors in the adjustment to industrial life among employees of a firm in south Nigeria. (Mimeographed report, Department of Social Anthropology, University of Edinburgh.)

WOBER, M. (1967*a*). Adapting Witkin's field independence theory to accommodate new information from Africa. *Br. J. Psychol.* **58**, 29–38.

WOBER, M. (1967*b*). Notes on administering psychological tests in Africa. *Bull. Br. psychol. Soc.* **20**, no. 68, 25–34.

Manuscript received 11 *March* 1969

Revised manuscript received 13 *July* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 114–122
With 1 text-figure
Printed in Great Britain

Self-Esteem and Sex as Factors Affecting Influenceability

By SVENN LINDSKOLD

Ohio University

AND JAMES T. TEDESCHI

State University of New York at Albany

A total of 60 fifth-graders (30 males, 30 females) were drawn from high and low scorers on the Coopersmith self-esteem (S-E) scale. In a modified Prisoner's Dilemma Game played for M & M candies, subjects were made the targets of either promises of extra reward for cooperating or threats of punishment for not cooperating which were sent by a simulated other player.

High S-E subjects rationally complied more to both threats and promises than did low S-E subjects, giving support to Cohen's (1959) hypothesis that high S-E is associated with avoidance or denial defences, while low S-E is associated with expressive or destructive defences. Subjects receiving promises were more open and truthful in indicating their compliance intentions than were threat-condition subjects; subjects in both message conditions were more cooperative overall than were no-message control subjects. A series of interactions of S-E with sex suggested a confounding effect of adopted sex-role as a possible explanation for the frequent failure to find with female subjects the direct relationship of low S-E to persuasibility or compliance that is typically found with males.

A recent review of the effects of fear-arousing communications (Higbee, 1969) cites a series of studies that consistently found that the effectiveness of warnings is directly related to self-esteem. The studies reviewed either asked subjects to indicate willingness to take some recommended preventive action to avoid harm or observed actual compliant actions taken. On the other hand, it has been found that self-esteem is negatively related to persuasibility (Hovland & Janis, 1959), particularly for males (Janis & Field, 1959; Gergen & Bauer, 1967). The persuasibility studies typically employed reports of opinion or preference change as dependent variables. The contradictory findings could be related either to the fact that individuals differing in self-esteem react in equal but opposite ways to negative and positive influence attempts or to the fact that with one set of findings behavioural compliance served as the primary measure, whereas, in the other case, expressions of attitudes constituted the dependent variable of interest. In the present experiment high and low self-esteem subjects were placed in a conflict-of-interests situation and received either threats or promises from a simulated opponent. The primary dependent variable was behavioural compliance. If the nature of influence is crucial to the compliant response, an interaction of self-esteem and type of influence should be found. If the nature of the dependent variable is the crucial factor, then self-esteem should be positively related to behavioural compliance irrespective of whether the influence message is negative or positive in nature.

Horai & Tedeschi (1969) modified the Prisoner's Dilemma Game (PDG) in order to study systematically the effects of threats upon behavioural compliance of target subjects. The PDG is a two-person, two-choice game which embodies the conflict between cooperation and competition. If both players simultaneously make the cooperative choice, they gain more than they do when both compete; however, if one cooperates while the other competes, the latter gains more than he would through joint cooperation and the former does worse than he would if both had competed. Hence, both stand to gain by cooperating, but there exists the constant temptation to exploit the opponent to do even better. Horai & Tedeschi had a simulated source, prior to certain specified PDG trials, send subjects explicit threats of punishment for non-cooperation on the next trial. The punishment was a penalty of additional points. The simulated source always competed himself on the threat trials; therefore, if the subject complied (cooperated), he always lost the maximum number of PDG points, but if he defied the threatener he could be penalized with an even larger loss. It was found that compliance was greater with increasing magnitude and probability of punishment. Lindskold *et al.* (1970) obtained comparable results with promises of additional rewards which were sent from a simulated source to preadolescent subjects. Not only does this experimental paradigm enlist the enthusiastic cooperation of subjects and hence apparently possesses a high degree of experimental realism (Aronson & Carlsmith, 1969), but it seems to possess a great deal of validity (Tedeschi, 1970).

In the present study, a 2 × 3 × 2 orthogonal design was employed. Subjects were separated into high and low self-esteem groups on the basis of pretest scores on the Coopersmith test of self-esteem and were assigned to three experimental conditions: (1) targets of threats; (2) targets of promises; and (3) a no-message unmodified PDG. Since the literature suggests sex differences, both male and female subjects were used.

METHOD

Subjects. A total of 60 fifth-graders (30 males, 30 females) attending the West Laboratory School in Coral Gables, Florida, served as subjects in the experiment.* Prior to the experiment all students in three fifth-grade classes completed the Coopersmith (1967) scale of self-esteem. Thirty-one subjects (16 males, 15 females) with scores of 29 or below ($\text{mean} = 23{\cdot}9$; $\text{range} = 7\text{–}29$) were assigned to the low S-E group, and 29 subjects (14 males, 15 females) with scores of 40 or above ($\text{mean} = 44{\cdot}3$; $\text{range} = 40\text{–}50$) were assigned to the high self-esteem group.† The self-esteem mean scores for the two groups were significantly different ($t = 20{\cdot}6$; $P < 0{\cdot}001$).

Procedure. After being summoned to the school cafeteria, the subjects were greeted by an adult male experimenter and led singly to one of three separated experimental areas. They were told that a game had been devised and that the experimenter was interested in how children of their age played it. It was explained that they would have a chance to win some

* The West Laboratory School is a public school; however, there is a long waiting list for admission and the students are not restricted to those residing in a specific geographic area. Many of the students are children of University of Miami faculty members. Approximately 82 per cent of the children used in the study were white, 18 per cent black. The blacks were rather evenly distributed across experimental groups.

† An error in subject selection following the S-E test resulted in four high S-E males and six low S-E males being assigned to the promise condition.

M & M candies; when asked, all subjects said they liked M & Ms.* Then they were loaned $200 in play money and were told that at the end of the game they would have to repay the $200, but whatever they had left over could be exchanged—one play dollar for one M & M.

A microphone, earphones and switch-box (all non-operative), which were placed on the experimenter's table, were pointed out to the subject. The experimenter explained that they were used to communicate with another, similarly equipped, adult at another school (the name of a neighbouring school was given if the subject asked) who was with another boy (girl) in the same grade. Repeated reference was made to the fact that the other subject was of the same sex. The subjects were then told that both players would be simultaneously turning towards or away from a designated wall. If they both turned towards the wall (cooperative choice), they would both win $4. If they both turned away (competitive choice), they would both lose $4. If one turned towards while the other turned away, the former lost $5 and the latter won $5. The four possible outcomes were carefully, and as concretely as possible, related to the choices of both players. The subjects were questioned regarding the outcome resulting from each pair of choices; and if they were unable to answer correctly, they were excluded from the study. Only one such child was excluded. The concrete rewards (M & Ms) emphasized an individualistic set—to win as much as possible for self, disregarding the other's outcome.

The sequence of events began with the subject facing a neutral wall, at right angles to the choices wall. The pay-off matrix, which the experimenter had referred to in explaining the outcomes, was posted on the neutral wall so that the subject had the pay-offs directly before him at the time he considered each strategy choice. The experimenter, who sat to the rear of the subject as he faced the wall, gave the signal, 'Ready', at which time the subject was to face squarely the neutral wall. At the signal ,'Go', the subject turned his entire body towards or away from the designated wall. The experimenter then reported the subject's choice into the dead microphone and gave the appearance of listening for the report of the simulated player's (SP) choice. The experimenter then provided feedback concerning the SP's response choice, stated the outcomes, and paid or received from the subject the appropriate amount. Fifty such trials were played with the SP following a programmed schedule which was 50 per cent cooperative.† The cooperative choices were placed randomly throughout the schedule.

Fourteen of the 50 trials were designated in advance as message trials. Posted on a card under the matrix was the promise message, 'If you face toward the wall next time I'll give you $10', *or* the threat message, 'If you don't face toward the wall the next time I will take $10 away from you', together with three replies which were available to the subject: 'I will face toward the wall the next time' (M1),'I will face away from the wall the next time' (M2) and 'I don't want to say what I will do next time' (M3). Subjects were informed that the other player could send them the promise (threat) message whenever he wanted; the subjects were required to reply with one of the three other messages, but they could not initiate communications.

On the message trials, the subject was told, while awaiting the 'go' signal, that the other subject had informed the other experimenter that he was sending the message. The subject's choice of a reply was obtained and reported into the microphone, and then the 'go' signal was given. In the promise condition, the pay-off for the outcome was made; and, then, if the subject had complied by cooperating, the experimenter announced that the SP wanted also to pay off the $10 bribe (as promised). In the threat condition, the pay-off for the outcome was made; and, then, if the subject had *not* complied, the experimenter announced that the SP was taking $10 from him (as threatened). Thus both promises and threats were 100 per cent credible. If no reward or punishment was due, the experimenter said nothing after the PDG pay-off was made and continued with PDG play until the next message trial. Seven of the 14 message trials were programmed as cooperations (facing the wall responses) by the SP.

---

* Sweet chocolates somewhat smaller than the British-made 'Smarties'.

† The use of a 50 per cent cooperative programmed schedule on the non-message trials was dictated by the fact that prior research with children, using the same procedure, had used a 50 per cent schedule; therefore comparisons were facilitated. Much PDG literature has shown that quite wide variations in the proportion of cooperations in a fixed programme of strategy choices have little differential effect on subjects' responding; hence the neutral, 50 per cent proportion was employed.

Compliance to either promises or threats was the least costly response overall for subjects. Because the SP was only 50 per cent cooperative on message trials, compliance to either message did often result in a loss of \$5 to the subject while the SP won \$5. However, the simulated source always did (100 per cent credibility) what he said he would do as regards bribes and punishments. Thus a subject in the promise condition might lose \$5 for compliance but would gain the \$10 bribe; the subject in the threat condition might lose \$5 for compliance but always avoided the costs for non-compliance—the negative game pay-off of \$4 and the punishment of \$10 (total costs of \$14).

The subjects were told that the other experimenter had a bank of money from which he paid the SP, or into which SP paid, in exactly the same manner as the subject did. The subjects did not appear to question the deception procedures. The four experimenters, who were randomized across conditions, ran several practice subjects to perfect their delivery of a believable patter.

At the conclusion of the 50 trials, the subject's money was counted, and he was told he would receive the M & Ms after all children had had an opportunity to play the game. Those who lost were told there would be a consolation prize, but they were not told what it was so that other subjects would not have this knowledge before their turn to play. The consolation prize was 25 M & Ms. The candy was distributed by school personnel at the conclusion of the approximately three weeks it took to finish the experiment.

RESULTS

Compliance behaviour. The no-message control group was necessarily excluded from the analysis of the replies and strategy choices made by the subjects in response to the promises and threats sent by the SP. Therefore a $2 \times 2 \times 2$ analysis of variance was performed on this set of dependent variables.

Only the factor of self-esteem had an effect on the primary measure of compliance—the frequency of cooperations made by the subjects on the 14 trials to which the messages referred. The high self-esteem subjects complied by cooperating an average of 11·26 times; the low self-esteem subjects complied 9·43 times ($F=4{\cdot}394$; d.f. $=1, 32$; $P<0{\cdot}044$). There were no other significant main effects or interactions on this dependent variable.

The subject's selection of a reply to the promise or threat gave an indication of attitudinal compliance on the one hand or of an effort to influence the strategy choice of the SP on the other. Unlike the decision to cooperate or to compete, the selection of a reply had no direct effect on the pay-off; therefore a cautiously compliant subject could have registered some ineffectual verbal defiance, or an exploitative subject could have announced cooperative intentions while planning to compete against the SP who would (the subject hoped) cooperate in the expectation of joining the subject in mutual cooperation. There were no significant main effects on the frequency of use of M1, the reply indicating the intention to cooperate; but there was an interaction of self-esteem by sex ($F=6{\cdot}327$; d.f. $=1, 32$; $P<0{\cdot}017$). Low self-esteem males and high self-esteem females used this message more than the other two groups; the means are given in Table 1. Only the three-way interaction of message-condition by self-esteem by sex was significant on the use of M2, the intent to compete ($F=4{\cdot}076$; d.f. $=1, 32$; $P<0{\cdot}052$). The interaction is depicted in Fig. 1 which shows that the high self-esteem males and the low self-esteem females departed to the greatest extent from the mean values; the high self-esteem males were particularly extreme in announcing defiance to the promise message.

Table 1. *Means illustrating the interactions of self-esteem with sex*

	High self-esteem		Low self-esteem	
Dependent variables	Males	Females	Males	Females
M1 frequency	6·22	8·30	8·73	5·50
M3 frequency	4·89	3·70	2·64	6·00
Truth with M2	0·568	0·786	0·869	0·678
Cooperation overall	0·479	0·536	0·542	0·471
Subject's pay-off matrix winnings	0·429	−2·400	−2·689	3·267
Cooperations on message trials*	11·11	11·40	10·64	8·10

* Not significant ($P<0{\cdot}105$).

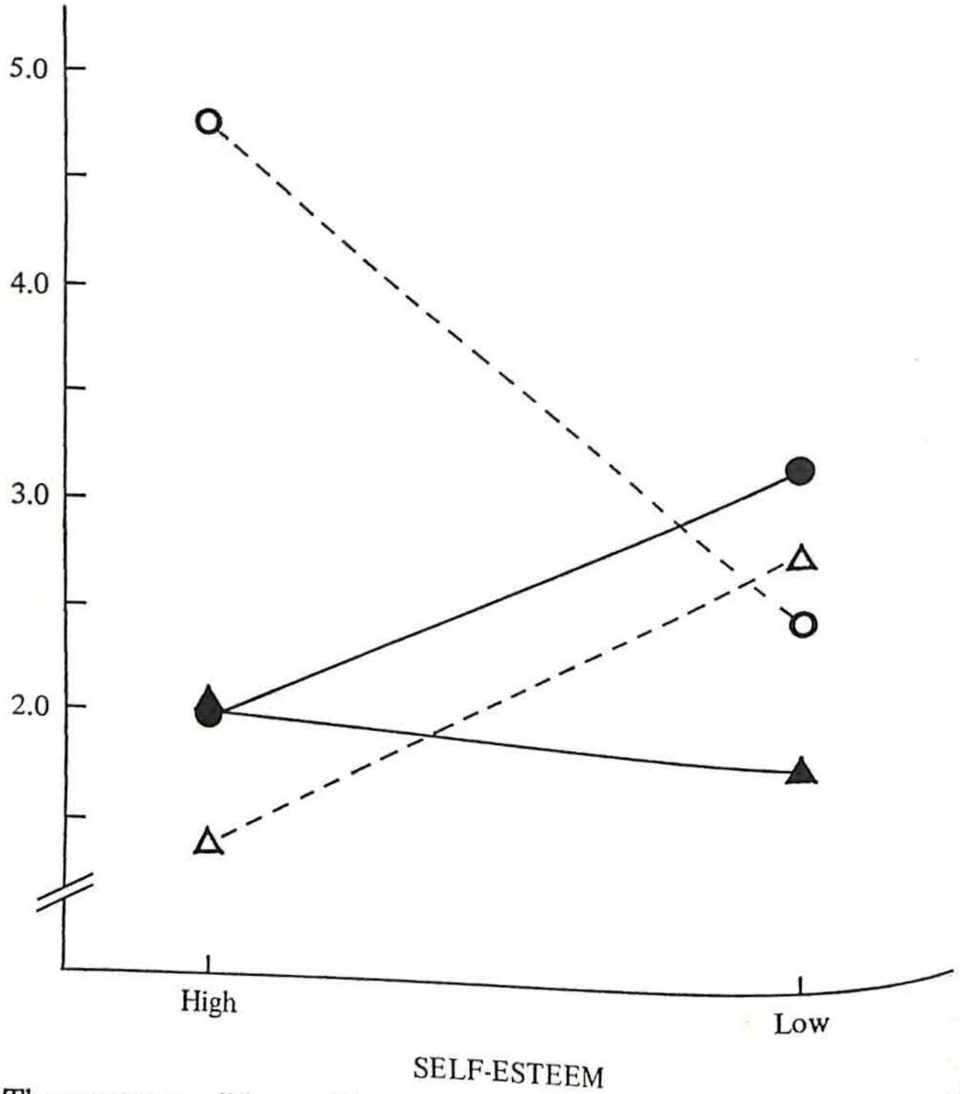


Fig. 1. The message-condition × self-esteem × sex interaction on the frequency of use of M2. Promise: ○----○, males; ●——●, females. Threat: △----△, males; ▲——▲, females.

There was a main effect of message-condition on the use of M3, the refusal to state intentions ($F=5{\cdot}746$; d.f. = 1, 32; $P<0{\cdot}023$). This message was used more frequently (mean = 5·40) in reply to the threat message than to the promise (mean = 3·10). There was also an interaction of self-esteem by sex ($F=4{\cdot}934$;

d.f. = 1, 32; $P < 0{\cdot}034$); the means given in Table 1 indicate that high self-esteem males and low self-esteem females used M3 more than the other two groups.

The strategy choices made by the subjects after using each of the three messages were also analysed. There were no significant effects of the proportion of cooperations made following the use of M1 or M3. There was, however, an effect of message-condition ($F = 6{\cdot}285$; d.f. = 1, 23; $P < 0{\cdot}020$) on the proportion of truthfulness (competing) after using M2 as a reply.* The subjects more truthfully (proportionately) announced that they would compete (non-comply) when they received the promise (mean = 0·864) than when they received threats (mean = 0·580). The interaction of self-esteem and sex was also significant on truthfulness after use of M2 ($F = 4{\cdot}181$; d.f. = 1, 23; $P < 0{\cdot}052$); the means shown in Table 1 indicate that low self-esteem males and high self-esteem females were more truthful.

Overall strategy and winnings. The no-message control group was included in the 3 × 2 × 2 analysis of variance of the subjects' proportion of cooperative strategy choices over all 50 trials and of the subjects' winnings derived from the pay-off matrix. The effects of the $10 rewards and punishments were so great on the final score accumulated by the subjects as to make the results trivial; therefore the winnings reported deal only with that which was gained from the 50 PDG trials alone.

There was a main effect of message-condition on cooperative proportion over all 50 trials (CP) ($F = 6{\cdot}131$; d.f. = 2, 48; $P < 0{\cdot}004$). Duncan multiple range tests indicated no difference between the promise (mean = 0·562) and the threat (mean = 0·541) conditions, while both differed ($P < 0{\cdot}01$) from the no-message control condition (mean = 0·421). Although the interaction of self-esteem by sex ($F = 3{\cdot}082$; d.f. = 1, 48; $P < 0{\cdot}086$) on CP fell slightly short of the accepted level of statistical significance, the result is reported because it confirmed the self-esteem by sex pattern found on many other dependent variables (see Table 1).

Message-condition also had an effect on winnings ($F = 4{\cdot}645$; d.f. = 2, 47; $P < 0{\cdot}014$). The range tests indicated that subjects won more (mean = 3·50) in the no-message condition than in the promise (mean = −4·00) condition ($P < 0{\cdot}01$). However, the threat condition (mean = −0·70) was not significantly different from either of the two others. Again, the interaction of self-esteem by sex was significant ($F = 4{\cdot}831$; d.f. = 1, 47; $P < 0{\cdot}033$). Table 1 shows that high self-esteem males and low self-esteem females won, while the other two groups suffered net losses.

The winnings of the SP reflected exactly the rank-order of subjects' winnings; the SP won 28·0 in the promise condition, 18·3 in the threat condition and −37·2 in the control condition. There was a significant effect of message-condition on the sum of winnings of 'both' players ($F = 6{\cdot}309$; d.f. = 2, 47; $P < 0{\cdot}004$). The joint winnings of subjects *and* SP were greater in the message conditions than in the control condition ($P < 0{\cdot}01$) but the former groups did not differ on this measure ($P > 0{\cdot}10$).

DISCUSSION

A positive relationship between self-esteem and behavioural compliance to threats and promises was the major finding of the present experiment. The failure to find an interaction of self-esteem with the negative or positive nature of the type of influence used by the simulated source indicates that previous contradictory findings may be attributed to differences in the dependent variables used. When the indicator taken focuses upon the behavioural alternative the individual says he prefers and/or upon the actual compliant behaviour taken, the function relating self-esteem to influenceability is positive. It still needs to be firmly established that when changes in attitude constitute the indicator, self-esteem and influenceability are negatively related.

---

* Nine subjects were excluded from this analysis because they made no use of message 2.

The compliant response was the most profitable one available to subjects even though the simulated player was exploitative on one-half of the message occasions; hence, if maximizing gains or minimizing costs within a situation is used as a criterion of rationality, the high self-esteem subjects acted on a more rational basis than did low self-esteem subjects since the former complied an average of more than 11 times out of 14 message occasions. That is, high self-esteem subjects were better able to reject the implications of being exploited than were the low self-esteem subjects. This result indicates that Cohen's (1959) association of high self-esteem with avoidance or denial defences and low self-esteem with the expressive, destructive defences is appropriate. As Scodel *et al.* (1959) concluded in an early PDG study exploring inter-subject communication, attempted collaboration can be ego-deflating; and the idea that the other player may be doing better than oneself is an assault on one's self-esteem. Kipnis (1968) found a tendency for high-impulsivity subjects to be more exploitative in a PDG in which the SP's choice was announced prior to the subject's choice; and Faucheux & Moscovici (1968), in a study which contrasted chronic and induced self-esteem, found a trend of greater exploitation by those subjects who had chronic low self-esteem.

Table 1 presents a striking series of replications of the self-esteem by sex interaction. The two middle columns of the table show high self-esteem females and low self-esteem males to be more openly and verbally compliant with their reply messages; less concealing; more constant when announcing defiance; and more cooperative overall, leading to smaller pay-offs—all conforming or approval-seeking behaviours. Beloff (1958) found that women of high self-esteem and men of low self-esteem were most conforming. On message-trial compliance, however, the high self-esteem males in the present study depart from the pattern to be, profitably, as compliant behaviourally as any other group. As Lynn (1962) has pointed out in his review of sex-role identification, males are better at problem-solving than females. This trend, expressed in greater winnings by males, is found generally in the PDG studies reporting sex differences (Gahagan *et al.*, 1969; Lindskold & Tedeschi, 1969; Tedeschi & Bonoma, 1969; Tedeschi *et al.*, 1969).

Perhaps because compliance was the more rational response in either case, there was no effect of message content on the principal compliance measure. The threat message did, however, promote greater frequency of use of the non-committal reply message (M3) and more lying with the defiant reply message (M2) prior to prudent compliance. Thus the data indicate that there was some verbal resistance to being threatened, perhaps a type of face-saving.

The finding that there was more cooperation overall when there were messages than in the no-message condition may be largely explained by the fact that the subjects in the message conditions were 74 per cent cooperative on the message trials. They were 48 per cent cooperative on the non-message trials as compared to the 42 per cent cooperations chosen by control subjects overall. There may have been some slight carry-over of the set to cooperate from the message trials to the non-message trials but probably not a significant one.

The slightly higher cooperation of the subjects in the promise condition might account for the results on winnings. Lindskold *et al.* (1970) and Lindskold & Tedeschi (1969) found the promise message had the effect, as contrasted with threats and no messages, of promoting greater overall cooperation. Since cooperation by subjects is generally inversely related to winnings in experiments in which the SP is 50 per cent cooperative, the finding that subjects in the promise condition won less than subjects in either of the other two conditions is consistent with prior results. Again, the winnings of the SP and the joint winnings of subjects *and* SP follow the pattern of cooperative responding by subjects. The SP won more when subjects cooperated more often. The winnings results, then, may be considered as artifactual concomitants of the degree of cooperativeness displayed by subjects across experimental conditions.

CONCLUSION

The present study indicates that, when self-esteem is involved, influence attempts directed towards opinion change may not be considered equivalent to influence attempts that relate to behaviour involving positive or negative outcomes to the subject. As Pepitone (1964) has suggested, a subject rewards himself in accord with his self-esteem; this self-reward is quite different from temporary attitudinal conformity of limited ego-relevance.

The difference between the sexes in the relations of self-esteem to persuasibility may be, as this study suggests, a function of less stereotyped sex-role adopted by females than by males. The evidence over many studies and many experimental tasks has consistently shown females, in general, to be more persuasible and conforming than males. But controlling for chronic self-esteem may also confound adopted sex-role, hence causing the resultant lack of clarity within sexes in the relationship of self-esteem to persuasibility and to behavioural compliance.

This research was supported by NIMH grant no. MH 17391-01. The authors wish to thank Dr James Gahagan, Noel Novinson, Thomas Bonoma, Charles Phillips, Brian McLean, and Peg Tedeschi for their help in carrying out the study, and Dr Thomas, the principal of the West Laboratory School, for his patient cooperation, while we caused minor disturbances to normal school routine. Most of all we appreciate the energetic and eager participation of the children who served as subjects.

REFERENCES

ARONSON, E. & CARLSMITH, J. M. (1969). Experimentation in social psychology. In G. Lindzey & E. Aronson (eds.), *Handbook of Social Psychology*, vol. 2 (2nd ed.). Reading, Mass.: Addison-Wesley.

BELOFF, H. (1958). Two forms of social conformity: acquiesence and conventionality. *J. abnorm. soc. Psychol.* **56**, 99–104.

COHEN, A. R. (1959). Some implications of self-esteem for social influence. In C. I. Hovland & I. L. Janis (eds.), *Personality and Persuasibility*. New Haven: Yale University Press.

COOPERSMITH, S. (1967). *The Antecedents of Self-esteem.* San Francisco: Freeman.

FAUCHEUX, C. & MOSCOVICI, S. (1968). Self-esteem and exploitative behaviour in a game against chance and nature. *J. Person. soc. Psychol.* **8**, 83–88.

GAHAGAN, J. P., LONG, H. & HORAI, J. (1969). Race of experimenter and reactions to threats by black preadolescents. *Proc. 77th Ann. Conv.* Washington, D.C.: American Psychological Association.

GERGEN, K. J. & BAUER, R. A. (1967). Interactive effects of self-esteem and task difficulty on social conformity. *J. Person. soc. Psychol.* **6**, 16–22.

HIGBEE, K. L. (1969). Fifteen years of fear arousal: research on threat appeals: 1953–1968. *Psychol. Bull.* **72**, 426–444.

HORAI, J. & TEDESCHI, J. T. (1969). Effects of credibility and magnitude of punishment on compliance to threats. *J. Person. soc. Psychol.* **12**, 164–169.

HOVLAND, C. I. & JANIS, I. L. (1959). *Personality and Persuasibility.* New Haven: Yale University Press.

JANIS, I. L. & FIELD, P. B. (1959). Sex differences and personality factors related to persuasibility. In C. I. Hovland & I. L. Janis (eds.), *Personality and Persuasibility*. New Haven: Yale University Press.

KIPNIS, D. (1968). Studies in character structure. *J. Person. soc. Psychol.* **8**, 217–227.

LINDSKOLD, S., CULLEN, P., GAHAGAN, J. & TEDESCHI, J. T. (1970). Developmental aspects of reaction to positive inducements. *Dev. Psychol.* (in press).

LINDSKOLD, S. & TEDESCHI, J. T. (1969). Effects of contingent promises on interpersonal conflict. (University of Miami, mimeographed manuscript.)

LYNN, D. B. (1962). Sex-role and parental identification. *Child Dev.* **33**, 555–564.

PEPITONE, A. (1964). *Attraction and Hostility*. New York: Atherton.

SCODEL, A., MINAS, J. S., RATOOSH, P. & LIPETZ, M. (1959). Some descriptive aspects of two-person games. *J. conflict Resol.* **3**, 114–120.

TEDESCHI, J. T. (1970). Threats and promises. In P. Swingle (ed.), *The Structure of Conflict*. New York: Academic Press.

TEDESCHI, J. T. & BONOMA, T. (1969). The effects of retaliation on the behavior of a threatening source. *Proc. 77th Ann. Conv.* Washington, D.C.: American Psychological Association.

TEDESCHI, J. T., FEUER, L. & BONOMA, T. (1969). Social desirability and reactions to promisor credibility. (University of Miami, mimeographed manuscript.)

Manuscript received 6 *March* 1970

Revised manuscript received 4 *June* 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 123–131
Printed in Great Britain

Ego-Involvement and Herzberg's Two-Factor Theory of Job Satisfaction: An Experimental Field Study

By TOBY D. WALL, GEOFFREY M. STEPHENSON
Department of Psychology, University of Nottingham

AND

COLIN SKIDMORE
Personnel Department, Scott-Bader Ltd., Wollaston, Northants.

Job applicants were interviewed using questions derived from Herzberg's (1959) interview schedule as part of their selection interview. In this highly ego-involved situation they responded in a manner predictable from Herzberg's two-factor theory of job attitudes. However, individually matched control subjects, asked the same questions in a less formal situation, gave reponses which failed to support Herzberg's theory. It is concluded that Herzberg's results are a product of ego-defensive processes within individuals.

The two-factor theory of job satisfaction (Herzberg *et al.*, 1959) continues to command considerable interest. Porter (1966) noted that it appeals particularly to personnel managers. Indeed, suggestions from the theory have been put into practice in the form of 'job enrichment' (Paul *et al.*, 1969).

Despite its popularity, and the large amount of research it has generated, the validity of the theory remains open to question. The central hypothesis of the two-factor theory is that the determinants of job satisfaction ('motivators') are qualitatively different from the determinants of job dissatisfaction ('hygiene factors'). Five factors stand out as determinants of job satisfaction: 'achievement', 'advancement', 'recognition', 'responsibility' and 'work itself'. The importance of these factors as perceived determinants of job *dissatisfaction* is very small. Herzberg *et al.* (1959) labelled the factors 'motivators' and argued that they are *intrinsic* to the performance of work. In contrast, the factors of 'company policy and administration', 'supervision', 'interpersonal relations' and 'work conditions' are said to be unimportant as determinants of job satisfaction but strongly related to job dissatisfaction. This set of factors Herzberg called 'hygiene factors' and argued that they are *extrinsic* to the performance of work: they represent the work *environment*. Thus a dichotomy is proposed whereby factors of one kind ('intrinsic') promote job satisfaction, whilst factors of a different nature ('extrinsic') determine job dissatisfaction. Herzberg claims that his findings are in direct opposition to the traditional idea that any job factor can cause both job satisfaction and job dissatisfaction.

Evidence bearing upon the two-factor theory, however, has been contradictory in nature. Myers (1964) obtained results supporting the theory in a population of both salaried and hourly paid employees; Schwartz *et al.* (1963) obtained similar results on a population of lower-level supervisors. Other investigations, using a variety of populations, have supported the two-factor theory (Halpern, 1966; Hahn, 1959; Herzberg, 1965). Evidence failing to support the two-factor theory has been equally abundant (Burke, 1966; Ewen, 1964; Ewen *et al.*, 1966; Graen, 1966; Malinovski & Barry, 1965; Hinrichs & Mischkind, 1967; Lahiri & Choudhiri, 1966).

Quinn & Kahn (1967) suggest that 'Those studies which best support the Herzberg theory are those methodologically most similar to Herzberg's original work'. Criticism of the Herzbergian methodology has taken several lines. House & Wigdor (1967) and Vroom (1964) point to the possible biases involved in content analysis. Hardin (1965) and Hinton (1968) offer evidence of the unreliability of the method used by Herzberg. Of especial importance is Vroom's criticism that 'It is . . . possible that obtained differences between stated sources of satisfaction and dissatisfaction stem from defensive processes within the individual respondent. Persons may be more likely to attribute the causes of satisfaction to their own achievements and accomplishments on the job. On the other hand, they may be more likely to attribute their dissatisfaction not to personal inadequacies or deficiencies, but to factors in the work environment: i.e. obstacles presented by company policies or supervision'. Vroom (1964) argued that this defensiveness could account for Herzberg's results. Kahn (1961) and House & Wigdor (1967) made similar criticisms. Wernimont (1966) compared 'intrinsic' and 'extrinsic' factors as determinants of job satisfaction, using a free-choice method similar to that used by Herzberg, and a forced-choice method as a control for defensiveness. In the free-choice situation the results more closely resembled those obtained by Herzberg than did the results from the forced-choice situation.

The present study. This study investigates Vroom's hypothesis that Herzberg's results can be attributed to defensive processes within individuals. It was predicted that *individuals interviewed in a highly formal and highly 'ego-involved' situation would be more likely to attribute their past highly 'satisfying' experiences to those factors which Herzberg called 'motivators', and attribute their past highly 'dissatisfying' experiences to those factors Herzberg labelled 'hygiene' factors, than would individuals interviewed in a less formal and less ego-involved situation.*

METHOD

To test this hypothesis, two contrasting groups of respondents were asked questions similar to those used in Herzberg's original study. The experimental group answered the questions during the course of a formal job selection interview at a given company. Acceptance or rejection of the respondent's application for the desired position within the company clearly depended upon his performance in the interview: it was in his own interests to show himself in the best possible light. The control group consisted of employees of the same company, who answered the same questions during the course of an informal interview. The interviewer in this situation attempted to ensure that the respondent was fully at ease. The employee knew that nothing depended upon his performance in the interview. The interview was confidential and the interviewer was known to be independent of the management of the company. It was stressed that the respondent should be as accurate as possible in his replies. In short, the individual had less to gain by responding in a socially acceptable manner.

The experimental group comprised 14 male respondents. The control group comprised a further 14 respondents who were *individually* matched as closely as possible for occupational level and age from a pool of 55 respondents.

The questions asked were as follows. 'Think of a time when you felt exceptionally good or exceptionally bad about your job, either your present job or any other job you have had. This time should be one which lasted for at least three weeks and may have lasted for any number of years: it should also be a period of time during which your overall feelings about your job were consistently good or bad. Can you do this? Then please answer the following questions. How long ago did this happen? Can you describe specifically what made the feelings begin? Can you tell me more precisely why you felt the way you did at the time? Now can you think of a time when you felt exceptionally good/bad [whichever not previously described]. [Questions repeated.]'

It should be noted that the questions asked are similar to those used by Herzberg; however, the identification of the attitude includes an instruction to the respondent to think of an overall, lasting attitude, conforming to Herzberg's definition of a 'long-range' feeling about the job. This was done as it was felt that lasting feelings would be more important than transient ones. The replies elicited from each individual were recorded on cards. From the replies the 'first-level' factors were coded blind by one of the authors, according to Herzberg's system. The factors were also coded by a second author and there was substantial agreement (94 per cent) concerning the number of factors elicited and (91 per cent) concerning the categories into which they would fall. There was perfect agreement over whether each factor was a 'motivator' or 'hygiene' factor.

RESULTS

Table 1 presents the results in a form comparable to that used by Herzberg *et al.* (1959, p. 72). The length of each bar represents the percentage frequency of occurrence of a given factor in relation to the number of sequences elicited. (Percentages total more than 100, since more than one factor can occur in each sequence.)

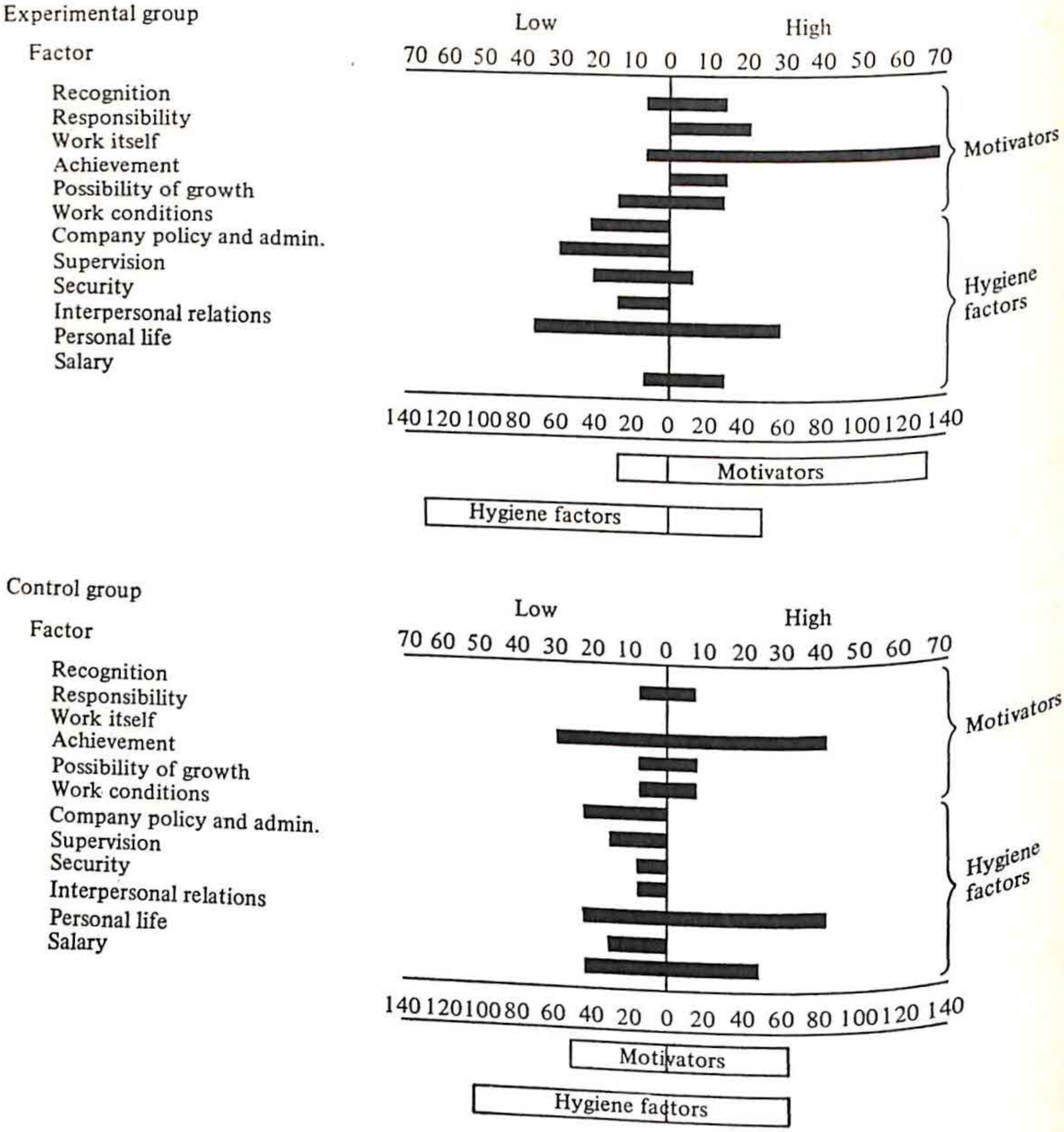
Inspection of this table reveals that the experimental group yields results similar to those reported by Herzberg *et al.*, while the control group does not. In the experimental group, 'motivators' (recognition, responsibility, etc.) are positively associated with high job attitudes (satisfaction), whilst the 'hygiene factors' (supervision, work conditions, etc.) occur mainly in connexion with low job attitudes (dissatisfaction). These trends are very much less marked in the control group. It is interesting to note that advancement, a 'motivator,' was not mentioned by any of the subjects.

An alternative method of presenting the results is more legitimate. In the previous method no allowance is made for the fact that some individuals yield several factors, while others yield only one. Verbose respondents may influence the results disproportionately. The alternative method used (Table 2) controls for the number of factors offered. Equal weight is attributed to each person's total contribution rather than to each factor. In practice this means that factors are given scores depending upon whether they were the *sole perceived reason* for the feelings reported, *one of two reasons*, *one of three*, *one of four* or *one of five* reasons. Thus where the factor was the sole reason it scored 100; where one of two, 50; one of three, $33\frac{1}{3}$; one of four, 25; and one of five, 20.

An inspection of the tables again reveals that the experimental group yields results similar to those reported by Herzberg, whereas the control group does not.

A statistical test of the hypothesis was carried out using a measure we have called the *Overall Herzbergian Score* (OHS). This score represents an individual's tendency to respond in the manner predicted by Herzberg's two-factor theory, i.e. the extent to which 'motivator' factors were perceived as determinants of past satisfaction and 'hygiene' factors were perceived as determinants of past dissatisfaction.

Table 1. *Percentage of each factor appearing in high and low job attitude sequences*

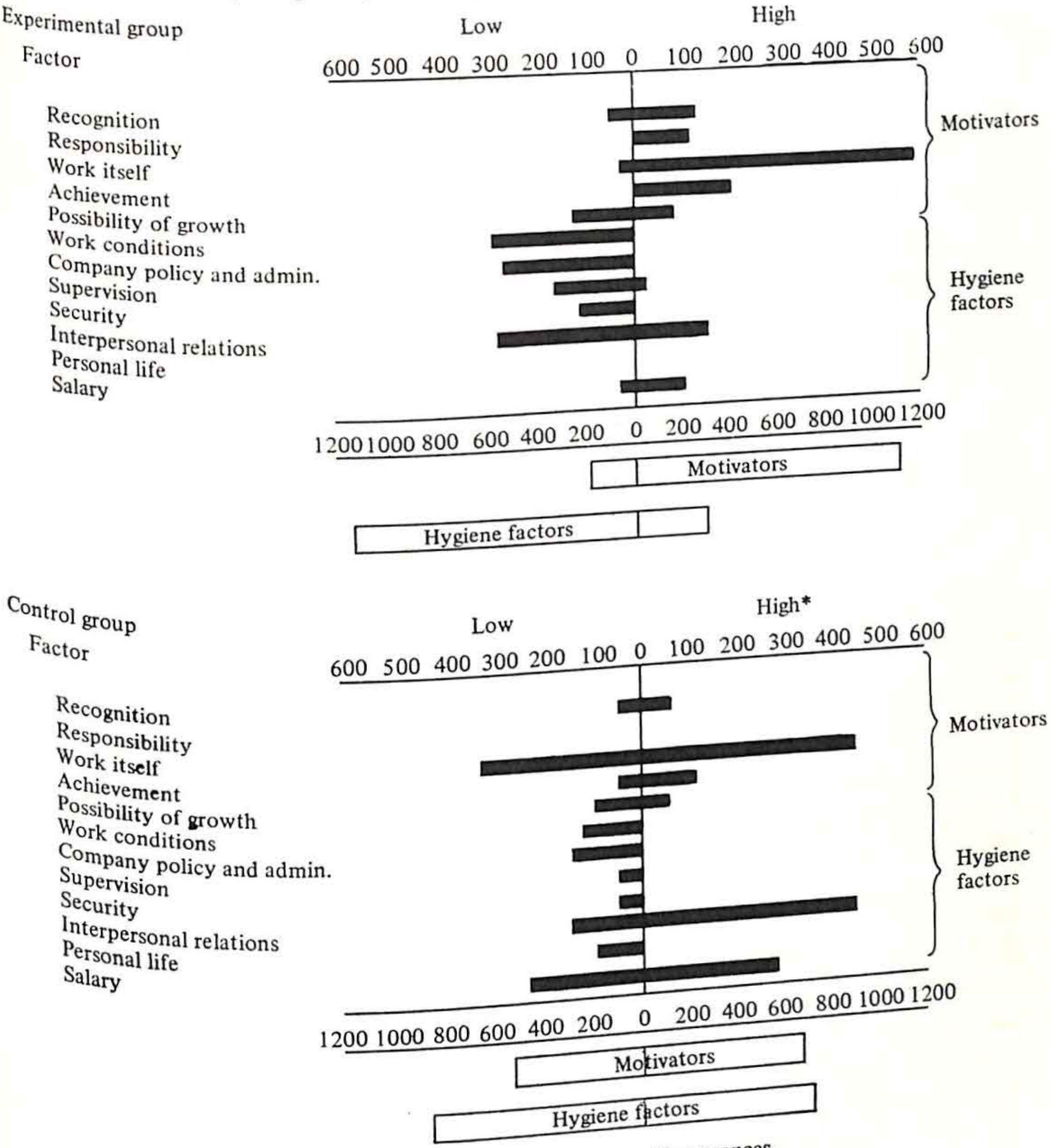


To take an example: Subject X gave three factors as reasons for satisfaction, two of which were motivators and one a hygiene. He also gave two reasons for dissatisfaction, both of which were hygiene factors. Where M = 'motivator' and H = 'hygiene factor', his results could be symbolized as follows: high job attitudes (satisfaction): MMH; low job attitudes (dissatisfaction): HH. Herzberg's two-factor theory is correct in 4/5 instances, giving this individual as OHS of 0·80.

Subject Y yielded the following results: HH for satisfaction and MH for dissatisfaction. His OHS score is 0·25, i.e. 1/4 correct.

Two similar analyses were used to establish whether responses to satisfaction or to dissatisfaction contributed differentially to the OHS results.

Table 2. *Sums of weighted factors appearing in high and low job attitude sequences*



* Adjusted for 12 sequences.

(*a*) *Satisfaction Herzbergian Score* (SHS). The two-factor theory predicts that motivators will contribute only to satisfaction. The SHS measures the proportion of motivators to hygienes in relation to satisfaction alone. To take an example: Subject X_1 gave three factors—HHM—as reasons for satisfaction. His SHS $= \frac{1}{3}$ or 0·33.

(*b*) *Dissatisfaction Herzbergian Score* (DHS). In relation to dissatisfaction the two-factor theory predicts that 'hygiene factors' will be the perceived determinants. To take an example: Subject Y_1 gave four factors—HHHM—as perceived determinants. The two-factor theory has made a prediction correctly in three out of four instances, yielding a DHS of 0·75.

Table 3 presents these three analyses in detail.

Table 3. *OHS, SHS and DHS scores of job selection interviewees (experimental group) and their matched controls (control group)*

	OHS			SHS			DHS		
Pair	Experimental	Control	Difference	Experimental	Control	Difference	Experimental	Control	Difference
1	1·00	0·50	+0·50	1·00	0·00	+1·00	1·00	1·00	0·00†
2	0·50	0·00	+0·50	1·00	0·00	+1·00	0·00	0·00	0·00†
3	0·80	1·00	−0·20	1·00	1·00	0·00†	0·75	1·00	−0·25
4	0·67	0·67	0·00†	0·50	1·00	−0·50	1·00	0·50	+0·50
5	0·75	0·67	+0·08	0·67	—	—*	1·00	0·67	+0·33
6	1·00	0·50	+0·50	1·00	1·00	0·00†	1·00	0·00	+1·00
7	0·67	0·00	+0·67	0·50	0·00	+0·50	1·00	0·00	+1·00
8	0·67	0·60	+0·07	0·50	0·33	+0·17	1·00	1·00	0·00†
9	0·80	0·75	+0·05	1·00	0·50	+0·50	0·67	1·00	−0·33
10	0·67	1·00	−0·33	1·00	1·00	0·00†	0·50	1·00	−0·50
11	1·00	0·50	+0·50	1·00	0·00	+1·00	1·00	1·00	0·00†
12	0·60	0·67	−0·07	0·50	0·00	+0·50	1·00	1·00	0·00†
13	0·67	0·50	+0·17	0·50	—	—*	1·00	0·50	+0·50
14	1·00	0·50	+0·50	1·00	1·00	0·00†	1·00	0·00	+1·00
Wilcoxon T value			$T=15{\cdot}5$ $P<0{\cdot}025$			$T=3{\cdot}5$ $P<0{\cdot}025$			$T=8{\cdot}5$ $P<0{\cdot}06$

* No reasons were given as respondents could not recall feelings of exceptional satisfaction.
† Tie.

Wilcoxon's matched-pairs signed-ranks test shows that the experimental group has significantly higher scores in all three cases. When individual SHS and DHS scores were averaged, $T=7{\cdot}5$, $P<0{\cdot}02$.

DISCUSSION

The field experiment reported in this paper supports Vroom's suggestion that Herzberg's results are an artifact of defensive processes within individual respondents.

The experimental group who were asked questions similar to those used in Herzberg's original study, during the course of a job selection interview, yielded results which were predictable from the two-factor theory. Each member of the experimental group was interviewed individually in a formal and highly ego-involved situation, where the impression the respondent gave of himself would have important implications for his future working life. It was clearly in the interest of each respondent in this group to present himself in the best possible light. In contrast, individually matched respondents in the control group, when asked the same questions, yielded results which were not compatible with the two-factor theory. The members of the control group were interviewed in an informal situation where nothing depended upon their performance in the interview; the control respondents knew the interviewer to be an independent agent and accepted that the interview was entirely confidential. Any need the control group members might have felt to give socially acceptable answers was thereby minimized, whereas this need was maximized for the experimental group. Tables 1–3 present the results analysed in different ways.

Table 1 presents the results in a manner directly comparable to Herzberg's results (1959). In the experimental group 'recognition', 'responsibility', 'work itself' and 'achievement' are all more often reported in connexion with 'high job attitudes' than they are with 'low job attitudes'; 'work conditions', 'company policy and administration', 'supervision', 'security' and 'interpersonal relations' are more often related to 'low job attitudes' than they are to 'high job attitudes'. In the control group this tendency is decidedly less marked. Comparing each factor across both conditions, every 'motivator' (except 'possibility of growth') in the experimental group is more strongly related to 'high job attitudes' and less strongly related to 'low job attitudes' than every motivator in the control group. Considering the 'hygiene factors': every 'hygiene factor' in the experimental group (except 'salary') is more strongly related to 'low job attitudes' and less strongly related to 'high job attitudes' (except 'supervision') than is every 'hygiene factor' in the control group.

Table 2 presents the results in a manner which controls for the varying numbers of factors related to job attitudes elicited from different individuals. The results lead more forcefully to the same conclusions as do the results of Table 1: that the experimental group display a relationship of 'motivator' and 'hygiene' factors to 'high' or 'low' job attitudes which are compatible with the two-factor theory, whilst the control group's results are incompatible with the theory.

Table 3 presents the results in terms of OHS, SHS and DHS scores. This is the most sensitive way of testing the difference between the results of the experimental and control groups, making use of the matched pairs. The results strongly substantiate Vroom's hypothesis that Herzberg's (1959) results could be interpreted as a consequence of defensive processes within individuals.

The number of respondents interviewed does not permit definite conclusions concerning the perceived determinants of 'high' and 'low job attitudes' to be made. However, it is interesting to note that the most important single factor was 'work itself'; this factor was also strongly affected by the situational manipulation of defensive processes within individuals.

The present authors suggest that Tables 2 and 3, which control for the different numbers of factors elicited for each respondent as perceived determinants of his attitude, present the evidence more appropriately than does Table 1. Herzberg's method of analysis (as in Table 1) gives equal weight to every factor, irrespective of whether it was perceived as a partial or total determinant of the job attitude being reported. For example, his results for high job attitudes are presented as the 'Percentage of Each First-level Factor Appearing in High and Low Job-Attitude Sequences' (1959, p. 60). There follows a statistical test upon the 'differences of totals between high and low' (1959, p. 72). This test could be meaningless without knowledge of the relative importance of each factor. If one accepts that a factor which is seen as one of several determinants of an attitude is not so important to that attitude as a factor which appears as the sole cause, then it is not possible to determine from Herzberg's results the extent to which they support his theory. The error of analysis made by Herzberg (1959) is repeated in a reanalysis of his findings by Hackman (1969) (cf. criticism by Lindsay *et al.*, 1967).

CONCLUSION

The results reported in this paper suggest that ego-defensive processes within individuals affect the responses they give to job attitude questions. This has strong implications for job attitude research in general and for the two-factor theory in particular.

Generalizations from any job attitude research depending solely upon evidence acquired in a job application situation (e.g. Jurgensen, 1948) must be viewed with extreme caution. Future work should make every effort either to minimize the effect of defensive processes or include in its design a means of assessing and partialling out their effects. In particular, we have shown that statements by employees which are in accord with Herzberg's two-factor theory of job motivation are an artifact of ego-defensive processes. In the light of this result and the increasing incidence of investigations which fail to support the theory, it is evident that the two-factor theory of employee motivation is unacceptable.

Financial support from the Social Science Research Council and from The Scott-Bader Commonwealth is gratefully acknowledged.

REFERENCES

BURKE, R. J. (1966). Are Herzberg's motivators and hygienes unidimensional? *J. appl. Psychol.* **50**, 317–321.

EWEN, R. B. (1964). Some determinants of job satisfaction: a study of the generality of Herzberg's theory. *J. appl. Psychol.* **48**, 161–163.

EWEN, R. B., SMITH, P. C., HULIN, C. L. & LOCKE, E. A. (1966). An empirical test of the Herzberg two-factor theory. *J. appl. Psychol.* **50**, 544–550.

GRAEN, G. B. (1966). Addendum to 'An empirical test of the Herzberg two-factor theory'. *J. appl. Psychol.* **50**, 551–555.

HACKMAN, R. C. (1969). *The Motivated Working Adult.* American Management Association Inc.

HAHN, C. (1959). Dimensions of job satisfaction and career motivation. (Unpublished manuscript; in House & Wigdor, 1967.)

HALPERN, G. (1966). Relative contributions of motivator and hygiene factors to overall job satisfaction. *J. appl. Psychol.* **50**, 198–200.

HARDIN, E. (1965). Perceived and actual changes in job satisfaction. *J. appl. Psychol.* **49**, 363–367.

HERZBERG, F. (1965). Job attitudes in the Soviet Union. *Personnel Psychol.* **18**, 248–252.

HERZBERG, F. (1966). *Work and the Nature of Man.* New York: World Publishing Company.

HERZBERG, F., MAUSNER, B. & SNYDERMAN, B. (1959). *The Motivation to Work.* New York: Wiley.

HINRICHS, J. R. & MISCHKIND, L. A. (1967). Empirical and theoretical limitations of the two-factor hypothesis of job satisfaction. *J. appl. Psychol.* **51**, 191–200..

HINTON, B. L. (1968). An empirical investigation of the Herzberg methodology and two-factor theory. *Org. Behav. & Human Perf.* **3**, 286–309.

House, R. J. & Wigdor, L. A. (1967). Herzberg's dual-factor theory of job satisfaction and motivation: a review of evidence and a criticism. *Personnel Psychol.* **20**, 369–389.

Jurgensen, C. E. (1948). What job applicants look for in a company. In M. L. Blum (ed.), *Experimental Industrial Psychology*. New York: Prentice-Hall, 1952.

Kahn, R. L. (1961). Job factors, attitudes and effects. (Review of *The Motivation to Work* by F. Herzberg *et al.*). *Contemp. Psychol.* **6**, 9–10.

Lahiri, D. K. & Choudhiri, P. K. (1966). Perceived importance of job factors by technical and non-technical employees. *Personnel Psychol.* **19**, 287–296.

Lindsay, C. A., Marks, E. & Gorlow, L. (1967). The Herzberg theory: a critique and reformulation. *J. appl. Psychol.* **51**, 330–339.

Malinovski, M. R. & Barry, J. R. (1965). Determinants of work attitudes. *J. appl. Psychol.* **49**, 446–451.

Myers, M. S. (1964). Who are your motivated workers? *Harv. busin. Rev.* Jan/Feb, 73–88.

Paul, W. J., Robertson, K. B. & Herzberg, F. (1969). Job enrichment pays off. *Harv. busin. Rev.*, March-April, 61–79.

Porter, L. W. (1966). Personnel management. *Ann. Rev. Psychol.* **17**, 395–422.

Quinn, R. P. & Kahn, R. L. (1967). Organizational psychology. *Ann. Rev. Psychol.* **18**, 437–466.

Schwartz, H. A., Jenusaitis, E. & Stark, M. A. (1963). Motivational factors among supervisors in the utility industry. *Personnel Psychol.* **16**, 45–53.

Vroom, V. H. (1964). *Work and Motivation*. New York: Wiley.

Wernimont, P. F. (1966). Intrinsic and extrinsic factors in job satisfaction. *J. appl. Psychol.* **50**, 41–50.

Manuscript received 25 *March* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 132–137
Printed in Great Britain

Cross-national Comparisons of Catholic–Protestant Creativity Differences

By CHARLES MARINO*

University of Edinburgh

It has been argued that personality factors associated with authoritarianism–dogmatism are antithetical to creativity. This study attempts to determine if these findings can be demonstrated in real or naturally occurring groups and if they are affected by socio-political contexts. The groups compared are matched samples of Roman Catholic and Protestant students from the United States, Northern Ireland, Eire and Scotland. Psychological evidence of relatively high authoritarianism in Roman Catholics and sociological indications of relatively low creative production by Roman Catholics lead to the general hypothesis that Catholic students will perform less well on mental ability tests of creativity factors.

A two-way analysis of variance design (country *v.* religion) is employed to test the hypotheses. The results indicate that there are large and statistically significant differences between Catholic and Protestant students in the U.S. and Northern Ireland. Catholic students in these countries evidenced less originality, ideational fluency and spontaneous flexibility than Protestants. There were generally no differences between the groups in Eire and Scotland. It is concluded that the antithesis of dogmatism–authoritarianism to creativity can be demonstrated in naturally occurring groups and that these differences are affected by the socio-political contexts in which the groups exist.

The personality characteristics associated with authoritarianism–dogmatism are antithetical to those associated with creativity (Barron, 1963; Hudson, 1968). Furthermore, it has been posited that social situations and institutions which foster pressures to conform, emphasize external evaluation or seek to impose convictions and values on individuals inhibit creativity (Rogers, 1959; Torrance, 1964; Moustakas, 1967). Most of the research concerned with social factors affecting creativity has dealt with aggregates of individuals. This study is an attempt to determine if these findings can be demonstrated in naturally occurring groups.

AUTHORITARIANISM, ROMAN CATHOLICISM AND CREATIVITY

There is considerable evidence indicating that religious conservatives in general and Roman Catholics in particular are relatively more authoritarian and dogmatic than other religious groups (Allport & Kramer, 1946; Adorno *et al.*, 1950; Argyle, 1958, Quin, 1965; Fox, 1965). Given the antithesis of authoritarianism–dogmatism to creativity, it seems logical to expect that the creativity of conservative religious groups (e.g. Roman Catholics) would be relatively low.

If we define scholarship, research or the attainment of intellectual eminence as important indicators of creativity, then the few studies comparing Catholics with other religious groups support this expectation. Knapp & Goodrich (1952) and Knapp & Greenbaum (1953) found that Catholics were disproportionately under-represented in their surveys of the origins of American scientists and scholars respectively. Lieberman (1960) came to similar conclusions about the intellectual contributions of Catholic thought and thinkers. (See also Brogan, 1941; Meng, 1957.) Argyle (1958, p. 96) states: 'It is possible that people taught to accept traditional ideas on authority do not easily become innovators or research workers.'

---

* Present address: Department of Psychology, York University, Toronto, Canada.

A second purpose of this research is to determine the generalizability of any differences which do occur. Clearly, naturally occurring groups are affected by the broader socio-cultural contexts within which they operate. There is some indication that *British* Catholic and Protestant students will exhibit hypothesized differences in creativity (Argyle, 1958; Scott, 1967; Hudson, 1968). Yet the extent of these differences is likely to be affected by socio-political factors, for example in Belfast as opposed to Edinburgh.

HYPOTHESES

The basic hypothesis of this study is that Catholic students will evidence less creativity than will Protestant students. The aspects of creativity which are tested herein are originality, fluency, and flexibility. *Originality* is associated with open-mindedness, independence of judgement and unconventionality (Barron, 1963). Therefore it is expected that a group demonstrably more closed-minded (Quin, 1965) and conventional (Allport *et al.*, 1948) will be less original than a group evidencing relatively less negative concomitance with this factor. *Fluency* is a measure of the free flow of ideas. One expects this ability to be antithetical to hypercriticality and dependence on external sources of evaluation (Guilford, 1967*b*; Rogers, 1959). Yet again, these variables are associated with authoritarianism, conventionalism and religious conservatism, each characteristic of Roman Catholics as a group. *Flexibility* is the ability to shift classes of response; as such it is indicative of a freedom from momentum in contrast to fluency, a freedom from inertia. Because authoritarianism is partially defined as the disposition to think in rigid categories (Brown, 1965) and because it is associated with emotional constriction, lack of spontaneity and repression of impulse (Barron, 1963), one expects a group exhibiting authoritarianism to be less flexible than a relatively antipathic group.

METHOD

Instruments. Originality was measured by the Uses Test (Guilford, 1967*a*). The subjects were asked to list as many different uses as they could think of for common objects (in this case a brick, pencil, paper clip, sheet of paper and a toothpick). Each response was subsequently weighted for statistical infrequency of occurrence and the score for originality is simply the sum of the weights of all the responses for any given subject. Fluency was measured in two ways: number of responses on the Uses Test (a measure of ideational fluency or the divergent production of semantic units) and through the Thurstone Word Fluency Test. The latter requires subjects to list as many words as possible beginning with a given letter. The number of words written represents a measure of word fluency or the divergent production of symbolic units. (In this case fluency is measured independently of meaning, hence 'symbolic' as opposed to 'semantic' production; cf. Guilford, 1967*a*). Flexibility was also measured in two ways: through the Uses Test which, when scored for the number of classes of response, yields a measure of spontaneous flexibility or the divergent production of semantic classes (Guilford, 1967*a*); and through the Word Association Test devised by Getzels & Jackson (1962). In the latter subjects are asked to list alternative uses for given words (e.g. Arm—gun, rocker, chair, of a person).

Subjects. The initial subject pool was composed of 500 subjects from 24 classes in 18 different schools located in the United States (Milwaukee, Wisconsin), Northern Ireland (Belfast), Eire (Dublin) and Scotland (Edinburgh). In each country two types of students, Catholic and Protestant, were tested. In every country the Catholics attended Catholic schools and the Protestants, non-Catholic schools. In Northern Ireland, Eire and Scotland the non-Catholic schools were Protestant affiliated, while in the United States they were public schools with no religious affiliation.

The subjects involved in the final matched samples were between the ages of 13 years 6 months and 14 years 6 months. The principal variable on which the subsamples were matched was IQ. Samples were further matched on the basis of socio-economic status (fathers' occupation) and sex. The final total of 200 subjects was composed of two samples of 25 subjects (Catholic and Protestant) in each of the four countries. The samples were principally composed of slightly above median IQ subjects (60th percentile; cf. Marino, 1968) from lower middle- and working-class backgrounds. The Catholic samples are all equally divided in terms of sex (13 females, 12 males), while the Protestant samples have a slightly greater proportion of females (14 females, 11 males).

Table 1. *F scores for two-way (country × religion) analyses of variance*†

	Originality	Ideational fluency	Flexibility	Word association	Word fluency
Country	0·39	3·03*	1·50	1·38	<1·00
Religion	21·00**	18·92**	14·26**	<1·00	1·88
Interaction	5·34**	2·68*	2·75*	<1·00	4·54**

* $P<0·05$. ** $P<0·01$.
† d.f.: country = 3; religion = 1; interaction = 3; error = 192.

Table 2. *Mean creativity scores*

	U.S.A.	Northern Ireland	Eire	Scotland
Originality				
Protestant	20·80	14·56	14·86	12·44
Catholic	8·32	9·60	10·40	12·51
Difference†	12·48***	4·96*	4·46*	−0·07
Ideational fluency				
Protestant	20·44	18·20	19·60	17·50
Catholic	13·60	13·88	18·90	14·88
Difference	6·84***	4·32**	0·70	2·62
Flexibility				
Protestant	15·72	13·88	14·10	12·36
Catholic	10·30	10·16	13·20	11·88
Difference	5·42***	3·72**	0·90	0·48
Word association				
Protestant	33·50	32·88	31·00	30·68
Catholic	34·50	29·20	32·60	30·10
Difference	−1·00	3·68	−1·60	0·58
Word fluency				
Protestant	24·13	29·48	25·50	27·32
Catholic	31·50	25·12	27·50	28·40
Difference†	−7·37**	4·36**	−2·00	−1·08

* $P<0·05$. ** $P<0·01$. *** $P<0·001$.
† Negative differences are in the non-hypothesized direction.

RESULTS

Two-way (country × religion) analyses of variance were conducted on scores from the tests of originality, ideational fluency, flexibility, word association, and word fluency. The results of these analyses are reported in Table 1. The mean scores for Catholics and Protestants in each country are reported in Table 2.

Originality, ideational fluency and flexibility scores of Catholic students in the United States and Northern Ireland are consistently lower than those for Protestants. With the exception of Eire (where Catholic students evidenced significantly less originality than Protestants) there were no significant differences between these groups in Eire and Scotland.

The Word Association Test did not distinguish between the groups in any country. The test of Word Fluency indicated significant differences in the U.S. and Northern Ireland but not in Eire or Scotland. The difference between Catholic and Protestant word fluency scores in the U.S. favoured the Catholics but was in the hypothesized direction for the Northern Irish groups.

DISCUSSION

Catholic students in the U.S. and Northern Ireland performed less well on tests of creativity than did matched samples of Protestant students. There were not differences on all tests, nor were there Catholic–Protestant differences in Ireland or Scotland. The antithesis of authoritarianism–dogmatism to creativity has been primarily argued in terms of individual differences. This study lends indirect support to this argument but, more importantly, it indicates the significance of social factors which affect creativity.

The antecedents of creative behaviour are probably best viewed as the interaction of personality and social factors. Neither an individual differences nor a social psychological approach is sufficient for an understanding of the phenomenon of creativity. Much of the research on creativity has proceeded from an individual differences perspective—including the many studies comparing groups whose *raison d'être* was simply and solely to compare personality characteristics in creatives *v.* non-creatives or creativity in individuals differing on various personality dimensions. There have been important exceptions to this primarily individual differences approach (e.g. Dentler & Mackler, 1964), but the social psychology of creativity is a largely neglected area of research. Clearly, if one is interested in the interaction of social and individual factors, studying aggregates of individuals and inferentially interpreting the results into theoretical statements about the effects of groups on individuals is less than direct.

This study was an attempt to translate a finding (the antithesis of authoritarianism–dogmatism to creativity) from an individual differences perspective to a social-psychological one. The analogues for this translation were real (i.e. naturally occurring) groups. Because the groups compared are 'real', they could be observed in their natural socio-cultural contexts and because of these contexts the translation from an individual differences to a social context was not literal. Two aspects of the results warrant further mention in this context: (*a*) the minimal differences in

Scotland, and (*b*) the anomalous performance of American Catholics on the Word Fluency test. Given the conservative religious context of Scottish education it is not surprising to find similarities between Catholic and Protestant students in Scotland. It is conservative religious groups in general that we expect to perform less well on tests of creativity.

The finding of significantly greater word fluency in American Catholics is anomalous in the context of this study. The symbolic as opposed to semantic (i.e. ideational fluency) production of divergent responses could be enhanced by a variety of uncontrolled variables in the present study. It seems possible, for example, that more traditional curricular emphases and teaching styles may facilitate symbolic fluency. Although further research would be required to determine the validity of this proposition, it has long been recognized that no one teaching method or curriculum design is best suited for all learning and types of learners.

The advice and encouragement of Dr Halla Beloff are gratefully acknowledged.

REFERENCES

ADORNO, T. W. *et al.* (1950). *The Authoritarian Personality*. New York: Harper.
ALLPORT, G. W., GILLESPIE, J. M. & YOUNG, J. (1948). The religion of the post-war college student. *J. Psychol.* **25**, 3–33.
ALLPORT, G. W. & KRAMER, B. M. (1946). Some roots of prejudice. *J. Psychol.* **22**, 9–39.
ARGYLE, M. (1958). *Religious Behaviour*. London: Routledge & Kegan Paul.
BARRON, F. (1963). *Creativity and Psychological Health*. New York: Van Nostrand.
BROGAN, D. W. (1941). *U.S.A.: An Outline of the Country, its People and Institutions*. London: Oxford University Press.
BROWN, R. (1965). *Social Psychology*. New York: Free Press.
DENTLER, R. A. & MACKLER, B. (1964). Originality: some social and personal determinants. *Behav. Sci.* **9**, 1–7.
FOX, J. T. (1965). Authoritarianism and the St. Ambrose College student. *Religious Educ.* **40**, 272–276.
GETZELS, J. W. & JACKSON, P. W. (1962). *Creativity and Intelligence*. New York: Wiley.
GUILFORD, J. P. (1967*a*). *The Nature of Human Intelligence*. New York: McGraw-Hill.
GUILFORD, J. P. (1967*b*). Some theoretical views of creativity. In H. Helson & W. Bevan (eds.), *Contemporary Approaches to Psychology*. Princeton, N.J.: Van Nostrand.
HUDSON, L. (1966). *Contrary Imaginations*. London: Methuen.
HUDSON, L. (1968). *Frames of Mind*. London: Methuen.
KNAPP, R. H. & GOODRICH, H. B. (1952). *Origins of American Scientists*. Chicago: University of Chicago Press.
KNAPP, R. H. & GREENBAUM, J. J. (1953). *The Younger American Scholar: His Collegiate Origins*. Chicago: University of Chicago Press.
LIEBERMAN, M. (1960). Parochial schools and public leadership. *Nat. Catholic educ. Ass. Bull.* **57**, 239–248.
MARINO, C. J. (1968). Creativity and conformity in school children as influenced by religious affiliation and type of school attended. (Unpublished Ph.D. thesis, University of Edinburgh.)
MENG, J. J. (1957). American thought: contributions of Catholic thought and thinkers. *Nat. Catholic educ. Ass. Bull.* **53**, 113–120.
MOUSTAKAS, C. (1967). *Creativity and Conformity*. London: Van Nostrand.
QUIN, P. V. (1965). Critical thinking and openmindedness in pupils from public and Catholic secondary schools. *J. soc. Psychol.* **66**, 23–30.
ROGERS, C. (1959). Toward a theory of creativity. In H. H. Anderson (ed.), *Creativity and its Cultivation*. New York: Harper & Row.
ROKEACH, M. (1960). *The Open and Closed Mind*. New York: Basic Books.

Scott, G. (1967). *The R.C.s: A Report on Roman Catholics in Britain Today*. London: Hutchinson.

Stein, M. I. (1953). Creativity and culture. *J. Psychol.* **36**, 311–322.

Torrance, E. P. (1964). Education and creativity. In C. W. Taylor (ed.), *Creativity: Progress and Potential*. New York: McGraw-Hill.

Weima, J. (1965). Authoritarianism, religious conservatism and sociocentric attitudes in Roman Catholic groups. *Hum. Relat.* **18**, 233–239.

Manuscript received 16 *April* 1970

Revised manuscript received 17 *July* 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 138–143
Printed in Great Britain

Psychometric Changes Associated with Response to Drug Treatment

By ALISTAIR E. PHILIP

Medical Research Council Unit for Epidemiological Studies in Psychiatry, Edinburgh University Department of Psychiatry, Edinburgh

Improved and unimproved depressed patients were compared on a number of psychological measures, each patient having been tested on three occasions. Analysis of variance showed that the Beck scale of depression, NSQ Anxiety and HDHQ Intropunitiveness showed differences between the groups and between occasions, there being significant interaction effects on the two former scales. NSQ Seriousness scores decreased over occasions, while NSQ Submissiveness, Tendermindedness and total score, along with the HOQ and HDHQ Extrapunitiveness, remained stable. These results are discussed in relation to other findings using the same or similar instruments. The need to distinguish between state and trait measures is discussed; the present study demonstrates how differences between psychiatric patients could be utilized in the study of function fluctuation.

It is commonly accepted that some aspects of the behaviour of individuals are stable over time while others fluctuate. When applied to psychiatric patients this trait-state distinction has resulted in attempts to identify personality variables and symptom variables which might be of use in following the course of illness (Adams & Foulds, 1962; Mayo, 1967). Some psychological measures, for example the Symptom Sign Inventory, or SSI (Foulds & Hope, 1968), clearly relate to states or conditions which are open to change with treatment or the passage of time. Other measures, such as the Hysteroid Obsessoid Questionnaire, or HOQ (Caine & Hope, 1967), relate to aspects of behaviour which are stable over time (Foulds, 1965). A third category, one in which a great many measures fall, relates to behaviour which has a stable component and a fluctuating component, on both of which will appear individual differences. Of this latter category, those measures which reflect mainly traits should be more stable over time than those in which state variance is present to a significant degree.

Philip (1969*b*) suggested that more information could be obtained from drug trials if the efficacy of treatment for each patient could be assessed. He presented a method for the statistical analysis of individual response to treatment, as assessed by a rating scale or other quantification of clinical judgement, which was a modification of Ferguson's non-parametric trend analysis of correlated observations (Ferguson, 1965). Sutherland *et al.* (1969) applied this method to drug trial data which had been analysed by conventional methods (Sutherland *et al.*, 1967). They concluded that the non-parametric trend analysis of ratings for individual patients was more informative than the conventional method. The present study reports on changes in some psychological measures in two groups of patients who participated in the above drug trial. A group of patients showing improvement and a group showing no improvement were compared on the Beck scale of depression (Beck *et al.*, 1961), the Hysteroid Obsessoid Questionnaire, the Extrapunitiveness and Intropunitiveness measures (Philip, 1969*a*) of the Hostility and Direction of Hostility Questionnaire, or HDHQ (Caine *et al.*, 1967) and the Neuroticism Scale Questionnaire, or NSQ (Scheier & Cattell, 1961).

It was expected that the following changes would occur over test occasions: (1) the Beck scores of both groups would show change, the improved group showing a greater reduction of scores than the unimproved; (2) scores on the HOQ would remain stable; (3) the HDHQ Intropunitiveness scores would decrease, especially in the improved group, while the Extrapunitiveness scores would change only to a minimal extent; (4) the NSQ, as a measure of 'neurotic tendency' and hence more of a trait than a state measure, should not change.

METHOD

Patients were admitted to the drug trial of Sutherland *et al.* (1967) if (*a*) a significant degree of depression was noted by the referring agent, admitting doctor or ward sister, and (*b*) a diagnosis of depression was made by one or both psychiatrists following interview. These female patients were rated on the Hamilton scale for depression (Hamilton, 1960) on admission, after two weeks' treatment and after four weeks' treatment. When these ratings were reduced to ranks and subjected to the procedure described by Philip (1969*b*) it was possible to produce two contrasting groups of patients. The first group comprised those women who showed a significant linear reduction in symptomatology, while the second comprised those who showed no significant change in their Hamilton ratings. Eighteen patients in each group had completed all the tests on all occasions and the ratio of trial drug to control drug patients, 4:5, was the same in both groups.

Analysis of variance using a repeated measurements design of Winer (1962) was carried out on each psychological variable. The design tested the significance of differences between the groups of patients, between test occasions, such occasions corresponding to the Hamilton rating occasions, and the interaction between occasions and groups.

RESULTS

The two groups of women have nearly identical mean Hamilton scale ratings on admission to hospital (improved group, 37·55 ± 8·38; unimproved, 37·50 ± 6·64). Table 1 shows the mean scores obtained for the various psychological measures.

Beck scale. There are very significant differences between the two groups of patients ($F = 15{\cdot}51$; d.f. = 1, 34; $P < 0{\cdot}001$) and between test occasions ($F = 18{\cdot}13$; d.f. = 2, 68; $P < 0{\cdot}001$). The interaction between these two variables is also significant ($F = 3{\cdot}36$, d.f. = 2, 68; $P < 0{\cdot}05$), although to a more modest degree. Both groups show a decrease in depression scores but the decrease is much more marked for the improved group. The unimproved patients have an initial mean score which is four points higher than that for the other group and a final score which is but two points lower than the improved group's starting point.

Hysteroid-Obsessoid Questionnaire. The HOQ remains stable over testing occasions and there is no difference between the groups, both scoring some five points lower (more obsessoid) than normal subjects.

Hostility and Direction of Hostility Questionnaire. The Extrapunitiveness measure does not change over time, nor is there a difference between the two groups. Intropunitiveness, on the other hand, shows differences between the groups ($F = 4{\cdot}33$;

Table 1. *Mean scores on tests*

	Improved ($n=18$)			Unimproved ($n=18$)			Group means			Occasion means				
	1st	2nd	3rd	1st	2nd	3rd	Improved	Unimproved		1st	2nd	3rd		G × O
Beck	20·7	11·3	6·7	24·9	22·8	18·7	12·9	22·1	$P<0.001$	22·8	17·1	12·7	$P<0\cdot001$	$P<0\cdot05$
HOQ	18·7	20·1	19·3	18·8	19·9	18·9	19·4	19·2	n.s.	18·8	20·0	19·1	n.s.	n.s.
HDHQ														
Extrapunitiveness	9·6	9·6	9·3	10·1	11·1	11·4	9·5	10·9	n.s.	9·9	10·4	10·4	n.s.	n.s.
Intropunitiveness	10·6	9·3	8·3	11·4	11·6	10·9	9·4	11·3	$P<0\cdot05$	11·0	10·4	9·6	$P<0\cdot05$	n.s.
NSQ														
Anxiety	8·7	8·0	7·4	9·2	9·1	9·2	8·0	9·2	$P<0\cdot01$	8·9	8·6	8·3	$P<0\cdot05$	$P<0\cdot05$
Submissiveness	8·2	8·4	8·3	8·1	8·6	8·2	8·3	8·3	n.s.	8·1	8·5	8·3	n.s.	n.s.
Seriousness	7·3	7·3	6·8	7·4	8·1	6·9	7·1	7·4	n.s.	7·4	7·7	6·9	$P<0\cdot05$	n.s.
Tendermindedness	5·2	5·9	5·7	5·6	5·6	5·6	5·6	5·6	n.s.	5·4	5·8	5·6	n.s.	n.s.
Total NSQ	8·7	8·6	8·3	9·0	9·1	8·6	8·5	8·9	n.s.	8·9	8·8	8·4	n.s.	n.s.

d.f. = 1, 34; $P < 0{\cdot}05$) and between occasions ($F = 3{\cdot}16$; d.f. = 2, 68; $P < 0{\cdot}05$). Patients who show no improvement are more intropunitive than improved patients on all three test occasions. There is a general drop in scores from initial testing onwards. Normal subjects achieve scores of 8·1 and 4·9 for Extrapunitiveness and Intropunitiveness respectively. It can be seen that on initial testing the present patients are slightly more extrapunitive and markedly more intropunitive than normal subjects.

Neuroticism Scale Questionnaire. The total Neuroticism score remains unchanged throughout the assessments as do the scores on Submissiveness and Tendermindedness. The mean scores on the first two of these scales are very high, a score of 5·5 being average in a normal population. The average scores on Tendermindedness support Kear-Colwell's (1965) finding that this factor does not discriminate normals and psychiatric patients in Scotland. The factor of Seriousness shows a significant difference between mean scores for occasions ($F = 3{\cdot}37$; d.f. = 2, 68; $P < 0{\cdot}05$). Improved and unimproved patients do not differ on this factor. The second-order factor of Anxiety shows results similar to those found for the Beck; there are significant differences between groups ($F = 9{\cdot}37$, d.f. = 1, 34; $P < 0.01$), between occasions ($F = 3{\cdot}19$; d.f. = 2, 68; $P < 0{\cdot}05$) and in the interaction between groups and occasions ($F = 3{\cdot}76$; d.f. = 2, 68; $P < 0{\cdot}05$). Unimproved patients are a little more anxious than the others at initial testing and this anxiety does not decrease. Patients who improve show a marked steady drop in anxiety scores over the three occasions, moving from a level which is unequivocally abnormal to one which is 'high normal'.

DISCUSSION

Comparing the mean scores obtained on initial testing by all the women taken together with the kind of scores which would be considered average in a normal sample, it can be seen that these patients are more obsessoid, more intropunitive and, extrapolating from the work of Metcalfe & Goldman (1965), express more depressed sentiments and feelings. Their expression of extrapunitive attitudes is well within normal limits, as are their scores on Tendermindedness in the NSQ; all the remaining NSQ scores are much higher than average. There is a slight tendency for the unimproved group to show more initial anxiety, intropunitiveness and depression but none of these tendencies can claim statistical respectability. More important than these is the fact that the two groups showed almost identical initial mean ratings on the Hamilton scale, changes on which determined the allocation of patients to the improved or unimproved categories.

The groups have behaved as predicted on the Beck scale, Intropunitiveness, HOQ, total NSQ, Submissiveness and Tendermindedness, showing change on the first two of these and no change on the others. It is true to say, however, that changes in the non-improved group are slight compared with changes in the improved. Such change behaviour is also true of Anxiety and Seriousness, the two NSQ scores which changed over occasions. On present evidence the NSQ is a test which has a mixture of trait and state measures. One of its measures, Tendermindedness, does not do what it is supposed to do; Kear-Colwell (1965) also found that this measure did not discriminate his patients from normals. The changes in symptomatology and Intropunitiveness, as well as the stability of the HOQ and Extrapunitiveness, are in accord with the findings of Mayo (1967). Metcalfe & Goldman (1965) present admission and discharge mean scores for the Beck which tally well with the figures reported here on improved patients. The unimproved patients have a final Beck score which is twice that of the discharge scores shown by these authors. The significant interaction effects on the Beck and Anxiety scales warrant some mention. Kerr *et al.* (1970) have reported that for reactive and endogenous depressives, the MPI Neuroticism scores at key illness of those patients who subsequently recovered were significantly lower than the scores of patients who remained unwell. A similar state of affairs seems to exist in this study, although the initial differences are not significantly different. The notion that depressed patients who present with very many complaints of being tense, upset and depressed at the time of admission are unlikely to be improved at follow up, can readily be put to the test.

Since Adcock (1965) has pointed to the similarity between Cattell's anxiety factor and Eysenck's neuroticism factor and Caine & Hope (1964) report a high correlation between the HOQ and Eysenck's extraversion factor, it is possible to compare some aspects of the present findings with those of Coppen & Metcalfe (1965). These authors report changes on both MPI dimensions, their patients showing a decrease on N and an increase on E so that their scores on recovery were similar to those of normals. The improved patients in the present study showed no change in HOQ scores, remaining more obsessoid than normals, while the reduction in their anxiety scores left them some way above the normal mean score. It would appear that changes in the level of emotional upset, be it measured by Eysenck's N or Cattell's Anxiety, do occur in depressed patients although the amount of change will vary according to the group studied. There are varied reports about change on the introversion–extraversion dimension. The present results and those of other workers (Adams & Foulds, 1962; Mayo, 1967) indicate that HOQ scores do not change over time for most patients, although melancholics may be an exception. Ingham's (1966) summary of studies using the MPI indicates that improvement in clinical state is accompanied by increased scores on the E scale. Differences in patients, time intervals and test measures contribute to this confused picture.

There are many test instruments used by psychologists which are a mixture of trait and state measures, the results produced by such crossbreeds being confusing and difficult to interpret. This study has shown that it is possible to identify measures which are essentially trait or state variables, using for this purpose demonstrable and replicable changes in the characteristics of psychiatric patients. Once a number of such measures have been identified they can be used as marker variables in examining the properties of other tests, this procedure leading not only to a better knowledge of tests in use but also to a greater awareness of the general problem of function fluctuation in personality characteristics.

Thanks are due to Drs M. S. and S. S. Sutherland who carried out the clinical assessments of patients and to the nursing staff at Kingseat Hospital, Newmachar, Aberdeenshire, who helped in the testing of the patients.

REFERENCES

ADAMS, A. & FOULDS, G. A. (1962). Depression and personality. *J. ment. Sci.* **108**, 474–486.

ADCOCK, C. J. (1965). A comparison of the concepts of Cattell and Eysenck. *Br. J. educ. Psychol.* **35**, 90–97.

BECK, A. T. *et al.* (1961). An inventory for measuring depression. *Archs gen. Psychiat.* **4**, 561–571.

CAINE, T. M., FOULDS, G. A. & HOPE, K. (1967). *Manual of the Hostility and Direction of Hostility Questionnaire* (*HDHQ*). London: University of London Press.

CAINE, T. M. & HOPE, K. (1964). Validation of the Maudsley Personality Inventory E scale. *Br. J. Psychol.* **55**, 447–452.

CAINE, T. M. & HOPE, K. (1967). *Manual of the Hysteroid-Obsessoid Questionnaire* (*HOQ*). London: University of London Press.

COPPEN, A. & METCALFE, M. (1965). Effect of a depressive illness on MPI scores. *Br. J. Psychiat.* **111**, 236–239.

FERGUSON, G. A. (1965). *Non-parametric Trend Analysis.* Montreal: McGill University Press.

FOULDS, G. A. (1965). *Personality and Personal Illness.* London: Tavistock.

FOULDS, G. A. & HOPE, K. (1968). *Manual of the Symptom Sign Inventory* (*SSI*). London: University of London Press.

HAMILTON, M. (1960). A rating scale for depression. *J. Neurol. Neurosurg. Psychiat.* **23**, 56–62.

INGHAM, J. G. (1966). Changes in MPI scores in neurotic patients: a three-year follow-up. *Br. J. Psychiat.* **112**, 931–939.

KEAR-COLWELL, J. J. (1965). Studies of the IPAT Neuroticism Scale Questionnaire (NSQ). *Br. J. soc. clin. Psychol.* **4**, 214–223.

KERR, T. A. *et al.* (1970). The relationship between the Maudsley Personality Inventory and the course of affective disorders. *Br. J. Psychiat.* **116**, 11–19.

MAYO, P. R. (1967). Some psychological changes associated with improvement in depression. *Br. J. soc. clin. Psychol.* **6**, 63–68.

METCALFE, M. & GOLDMAN, E. (1965). Validation of an inventory for measuring depression. *Br. J. Psychiat.* **111**, 240–242.

PHILIP, A. E. (1969*a*). The development and use of the Hostility and Direction of Hostility Questionnaire. *J. psychosom. Res.* **13**, 283–287.

PHILIP, A. E. (1969*b*). A method for analysing assessments of symptom change. *Br. J. Psychiat.* **115**, 1379–1382.

SCHEIER, I. H. & CATTELL, R. B. (1961). *Handbook for the Neuroticism Scale Questionnaire.* Champaign, Ill.: I.P.A.T.

SUTHERLAND, M. S., SUTHERLAND, S. S. & PHILIP, A. E. (1967). Depressive illness: comparison of effects of pramindole (WY3263) and imipramine. *Clin. trials J.* **4**, 857–860.

SUTHERLAND, S. S., PHILIP, A. E. & SUTHERLAND, M. S. (1969). The response to treatment of individual patients in a drug trial. *Br. J. Psychiat.* **115**, 1383–1385.

WINER, B. J. (1962). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

Manuscript received 15 *December* 1969

Revised manuscript received 16 *February* 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 144–151
Printed in Great Britain

Characteristics and Validity of the Grid Test of Thought Disorder

By D. BANNISTER
Bexley Hospital

FAY FRANSELLA*
St George's Hospital

AND JOYCE AGNEW
Crichton Royal Hospital

A sample of 316 psychiatric admissions were given the Grid Test of Thought Disorder (Bannister & Fransella, 1966, 1967). Test indices were shown to relate to case note judgements and marginally to prognostic data of the 'condition on discharge' type. The relationship between grid scores and diagnostic category was examined. A measure of the degree of abnormality of patterning of construct relationships (Social Deviation) was derived from the grid and shown to relate to structural thought disorder, sex and presence of a precipitating factor for the 'illness'. The process implications of the findings are discussed.

Two assumptions underlie the Grid Test of Thought Disorder. The first is that the grid is an effective general measure of the relationships between an individual's constructs. This contention has been examined extensively in Kelly (1955, 1969) and Bannister & Mair (1968). The second assumption is that persons clinically designated as thought-disordered schizophrenics manifest very weak relationships between their constructs and that such relationships as are manifest are unstable. The implications of this argument are examined in relation to experiments reported in Bannister (1960, 1962, 1963, 1965), Bannister & Salmon (1966), Al-Issa & Robertson (1964), Foulds *et al.* (1967), Presly (1969), Romney (1969), McPherson (1969), McPherson *et al.* (1970) and Bannister & Fransella (1971).

In Bannister & Fransella (1966, 1967) previous research forms of the thought-disorder grid were standardized in order to produce a clinically usable test of thought disorder. The present study was designed to reassess the validity of this measure, to throw light on some of its characteristics and to present an initial exploration of a further measure (Social Deviation) which can be derived from grid tests.

In conventional cross-validation studies criterion measures are usually solicited for the study and made to 'Sunday best' requirements. In this study the data were mostly 'soft', deriving from case notes. Manifest errors included confusion between disorders of content (delusions) and structural or formal thought disorder. However, case note judgements were accepted without 'revision' and if, additionally, the degree of interjudge unreliability in clinical assessments is borne in mind, it will be recognized as a relatively harsh context for cross-validation.

---

* Now at Bexley Hospital, Kent.

METHOD

Population

The population used for the study was a sample of 316 psychiatric admissions, aged between 17 and 60 and of Mill Hill IQ 80+, drawn from three different, large NHS hospitals (*n* is less than 316 in some analyses because not all measures were available for all cases). Some of the characteristics of the total population (144 men, 172 women) are shown in Table 1.

Table 1. *Characteristics of total sample*

	Mean	S.D.
Consistency	0·66	0·43
Intensity	1084	499
IQ (Mill Hill Scale)	103	13·4
Age (years)	37·9	11·5
Length of stay (weeks)	10·5	13·8

(*n* is 316 for all measures)

Consistency (the degree to which the *pattern* of intercorrelations between constructs is maintained from first to second grid) and Intensity (a crude measure of the total *amount* of interrelationship between constructs in both grids) are the two measures of thought disorder so far derived from the grid test. The recommended cut-off points below which thought disorder is suspected are 0·49 for Consistency and 1000 for Intensity. There was no significant association between age, sex or verbal intelligence on the one hand, and either Consistency or Intensity on the other. This confirms findings in the original standardization study and in related studies.

Cross-validation

Without knowledge of grid test results, 32 of the patients had been judged clinically thought-disordered and 168 judged not thought-disordered by the responsible psychiatrists. (In the remaining cases clinical judgement and grid results might have been contaminated.) Using the 'uncontaminated' population, the ability of the test to relate to the clinical judgement of thought disorder was again assessed and results for each measure separately are shown in Table 2.

Table 2. *Relationship of Intensity and Consistency to the clinical judgement of 'thought-disordered' for all patients for whom a clinical decision, uncontaminated by test results, was made*

	Clinically thought-disordered ($n=32$)		Clinically not thought-disordered ($n=169$)			*P*
	Mean	S.D.	Mean	S.D.	*t*	(one-tailed)
Intensity	912	398	1117	499	2·191	<0·025
Consistency	0·52	0·43	0·68	0·41	2·92	<0·005

When the two measures were combined, and the usual cut-off points used, a chi square of the relationship between the clinical and the grid decision of thought disorder yielded a chi-square value of 12·261 (d.f. = 1; $P < 0{\cdot}001$).

Table 3. *2 × 2 table of 'grid' and 'clinical' categorization of thought disorder*

		Clinically thought-disordered	
		Yes	No
Grid thought-disordered (< 0·49 Consistency < 1000 Intensity)	Yes	15	31
	No	17	137

Bearing in mind that these samples were not selected to be extreme groups, but were a randomly collected group of admissions, the results of the study support the utility of the grid test as a measure of thought disorder, *using clinical criteria.*

Grid test scores in relation to diagnosis

As in previous studies, grid scores of this kind seem not only to distinguish thought-disordered schizophrenics as a group, but also to scatter out other diagnostic groups in such a way that significant differences occur. Consistency and Intensity scores for five diagnostic categories are shown in Table 4.

Table 4. *Diagnostic group scores on Consistency and Intensity*
(Patients with disputed or uncertain diagnoses omitted.)

Diagnostic group	*n*	Consistency		Intensity	
		Mean	S.D.	Mean	S.D.
Neurotics	38	0·77	0·41	1420	552
Alcoholics	33	0·72	0·32	1177	428
Depressives	70	0·65	0·40	1051	541
Schizophrenics	88	0·59	0·48	983	427
Organics	6	0·28	0·23	736	287

The schizophrenics are the total mixed group of thought-disordered, not thought-disordered and 'no decision made' schizophrenics and differences between this group and other diagnostic categories are expected because it contains thought-disordered patients. However, other significant differences occur. Specifically, neurotics are significantly more consistent than depressives ($P < 0{\cdot}01$) and organics ($P < 0{\cdot}001$). Alcoholics are significantly more consistent than organics ($P < 0{\cdot}001$) and depressives also more consistent than organics ($P < 0{\cdot}001$). Neurotics have significantly higher Intensity scores than alcoholics ($P < 0{\cdot}05$), than depressives ($P < 0{\cdot}001$) and organics ($P < 0{\cdot}01$). Alcoholics have significantly higher Intensity than organics ($P < 0{\cdot}05$). All these tests are two-tailed.

Only the very small group of organics is lower in Intensity and Consistency than thought-disordered schizophrenics and, as was indicated in the standardizing study (Bannister & Fransella, 1966), it may be difficult to discriminate severely brain-damaged patients (particularly those with diffuse brain damage) from thought-disordered schizophrenics, using results on this test alone. More extensive grid investigation of organics is clearly required. At the opposite end of the scale, neurotics in this sample show high levels of Intensity and Consistency which separate them from other groups and there are occasional significant intervals in the mean scores of other diagnostic categories. It seems possible that Intensity and Consistency, as structural features of construct systems, relate to various diagnostic features but, while such relationships might have some research interests and implications of their own, the test in its present form is designed solely to identify thought disorder.

First and second grid

It has been observed clinically that subjects tend to increase their Intensity scores from the first to the second grid and this was confirmed at a highly significant level in the present study. The sample of 316 patients had a mean Intensity score of 471 (S.D. 239) for grid I and a mean of 615 (S.D. 299) for grid II: the difference is significant at the $P<0{\cdot}001$ level. This suggests that the articulation of construing necessary to complete a grid initiates a tightening process in itself. However, subjects who are not thought-disordered gain about 30 per cent Intensity as between the two grids, while thought-disordered subjects (in test terms) gain about 25 per cent as between the two grids, which may be taken to suggest that even at very low levels of structure, tightening remains a test (and maybe a therapeutic) possibility. Although Consistency and Intensity are mathematically independent measures, their positive correlation for this sample was 0·51. This is in line with previously found correlations (for *psychiatric* groups but not for normals) in Bannister & Fransella (1966).

Grid test scores and prognosis

Obviously, many aspects of psychiatric conditions unmeasured by the grid will affect the prognosis of a patient, but it could be argued that all the patients who have very low scores on the grid test might have a dubious prognosis in that thought disorder is deemed to represent a major breakdown in the integration of the person. Various aspects of prognosis, specifically length of current stay in hospital, number of previous admissions, number of admissions in the year following testing, total number of admissions up to one year after testing and condition on discharge were examined in relation to both Intensity and Consistency. Two aspects of prognosis were found to relate significantly in the predicted direction. The 128 patients who were rated as being in 'good' condition on discharge had a mean Intensity of 1097 (S.D. 497), while the 27 rated as being in a 'poor' condition at time of discharge had a mean Intensity of 880 (S.D. 317), these groups differing significantly ($P<0{\cdot}025$, one-tailed). Those rated 'fair' on discharge were disregarded in this calculation. The mean number of previous admissions for the group of 73 patients who were thought-disordered in grid terms was 1·67 (S.D. 99) and for patients not thought-disordered in grid terms ($n = 242$) the mean was 1·16 (S.D. 93). These means differed significantly ($P < 0·05$, one-tailed). No other aspects of prognosis related significantly to Intensity and Consistency. These results add only marginally to evidence of validity for the grid test but are of a kind worth seeking in that they are evidence different from, and independent of, the *psychiatrist's* judgement of 'thought-disordered'.

The Social Deviation score

Both Consistency and Intensity are essentially structural measures in that they do not relate to the particular *pattern* of relationships between constructs (the content of construing) but only to the presence and stability of the organization of constructs as such. In two previous studies (Bannister, 1960, 1962) the problem of the content or pattern of construing was tentatively examined, using in the first the raw score discrepancy between each individual's grid scores and the mean pattern for the total sample. In the second study a correlational measure which partialled out the effect of Intensity differences was used. In the first study schizophrenics who were not thought-disordered appeared to be more deviant in their pattern than other groups, and in the second study, using the correlational measure, thought-disordered schizophrenics appeared to be more deviant in whatever pattern they had managed to retain. In the present study this issue of content, i.e. the question of which constructs relate to which, positively or negatively, was tentatively explored.

As a crude measure of Social Deviation the following score was devised. Each of the 15 relationships between pairs of the six constructs of each grid has a modal direction, positive or negative, for any normal population, this being (for the constructs in the order in which they are used in the test) − − + − + + − + − − + − − + −. A simple count was taken, for each individual, of how often his correlation for each pair of the 30 pairs of constructs (15 for each grid) deviated in direction (positive or negative) from the norm. Zero correlations were ignored and a deviant correlation was counted only if it was in the opposite direction from the usual and if rho was 0·5 or greater. A rho of 0·5 was chosen because that for thought-disordered subjects is of this order. The aim here was to partial out the effect of differences in Intensity and this appears to have been achieved in that, for the 316 subjects, there was zero correlation between the Social Deviation score and Intensity.

If a simple count of differences in sign had been made, this would have maximized the Social Deviation score of thought-disordered schizophrenics who tend to have many correlations drifting, by chance, either side of zero. Such a form of Social Deviation score (as was found on checking in this study) tends to be too much a function of the Intensity score. For this reason the present scoring procedure (which partials out Intensity) was used.

The mean for the whole population of 316 for the Social Deviation score was 1·48 (S.D. 2·35). Working on the basis of the Bannister (1962) results, and arguing from the general premise that a distorted construct system is a bus stop on the way to thought disorder (albeit a bus stop at which some may get off and remain), it was predicted that the group of schizophrenics (which includes thought-disordered and not thought-disordered) would be significantly more socially deviant than other diagnostic groups. In fact, they were significantly more socially deviant than depressives. The 88 schizophrenics had a mean of 1·85 (S.D. 2·49) and 70 depressives had a mean of 1·14 (S.D. 2·18; $P < 0{\cdot}05$, one-tailed). The small sample of six organics again appeared socially deviant as well as weak in structure, having a mean of 3·17 (S.D. 3·12), which made them significantly more deviant than depressives ($P < 0{\cdot}05$, two-tailed).

The prediction of greater social deviance was then applied to the more select group of 32 *clinically* thought-disordered schizophrenics. They were found to have a mean Social Deviation score of 2·09 (S.D. 2·64), differing significantly from the 168 clinically not thought-disordered subjects with a mean of 1·33 (S.D. 2·25; $P < 0{\cdot}05$, one-tailed).

For thought-disordered subjects defined in grid test terms ($n = 73$) the mean was 2·11 (S.D. 2·09), compared with those not thought-disordered in grid terms ($n = 242$), whose mean was 1·30 (S.D. 2·39; $P < 0{\cdot}01$, two-tailed). The correlation between the Social Deviation score and Consistency was $-0{\cdot}21$ ($n = 316$), which is significant ($P < 0{\cdot}01$, one-tailed). The data support the general argument that a breakdown in structure may be initiated by deviant patterning. Ideas become odd before they become chaotic and lose their linkages.

One minor finding of some interest is that 14 patients of assorted diagnoses, who were described as having a clear precipitating factor for their illness, had a mean Social Deviation score of 2·50 (S.D. 2·50), while those 12 patients who were described as clearly having *no* precipitating factor had a mean of 0·33 (S.D. 1·16), giving a significant difference ($P < 0{\cdot}02$, two-tailed). We may speculate that the 'precipitating factor' here may refer to some single event, traumatic enough to require a special and peculiar 'adaptive' patterning of his constructs on the part of the confronted person.

The Social Deviation score, as here calculated, does not correlate significantly with either age or intelligence. However, there was a difference between the sexes (males: mean 1·86, S.D. 2·69; females: mean 1·17, S.D. 1·97, $P < 0{\cdot}02$, two-tailed), with females being less socially deviant. This finding can be pondered both for its 'social psychology' implications and for a possible relationship to the lower incidence of schizophrenia in women.

Clearly the Social Deviation score, as derived from this standard form of thought-disorder grid, is not usable as any kind of clinical measure. Firstly, its apparent correlates are few, although, such as they are, they have a bearing on the general argument as to the nature of the process whereby people become thought-disordered. More importantly, the test was designed specifically to maximize *structural* differences and, not surprisingly, it is not well designed to examine content differences. Thus the sample of constructs is probably too small in number to elicit any very reliable picture of the *pattern* of construct relationships for a subject in normative terms. Secondly, the constructs were selected so as all to relate highly positively or highly negatively for most subjects, thus maximizing the 'normal' Intensity score and helping to discriminate thought-disordered schizophrenics. Any grid test designed to examine the degree of social deviancy of *content* would need to include constructs which, for a normal population, had near zero relationships so that deviations which took the form of linking constructs, which were usually orthogonal, could also be made manifest.

However, the findings have two aspects of some importance. Firstly, they lean towards the general argument that thought disorder is an end point in a process in which, in the early stages, there may be violent changes in the usual pattern of construct relationships. This would tally with the results of the serial invalidation studies reported in Bannister (1963, 1965) which showed considerable changes in construct patterning as the *primary* reaction to the experience of being repeatedly wrong. Secondly, it suggests that the grid may be usable for the operational definition of 'social deviation' in construct patterning. The general distinction made in construct theory terms between structure and content requires careful handling since the two are related both operationally and theoretically (cf. Adams-Webber, 1969), but might well prove a profitable research line both in elucidating the characteristics of so-called abnormal conditions and in specifying normal processes of construct change.

CONCLUSIONS

The study has provided further validational evidence for the usefulness of the Grid Test of Thought Disorder in terms of its correlation with the clinical judgement of thought disorder. It has provided scant but initial evidence, independent of the clinical judgement of thought disorder, of the meaningfulness of grid test results in this field. Thirdly, it suggests a possible extension of grid-based construct analyses in the direction of 'social content' exploration.

The first author acknowledges the support of the Medical Research Council and all the authors appreciate the help given by the staffs of Bexley, Springfield and Crichton Royal Hospitals, in particular the following: Margaret Steinhard of Bexley, Cecile Gourney, William Price and Sue Mumford of Springfield and Dr A. C. Tait of the Crichton Royal.

REFERENCES

ADAMS-WEBBER, J. R. (1970). An analysis of the discriminant validity of several repertory grid indices. *Br. J. Psychol.* **61**, 83–90.

AL-ISSA, I. & ROBERTSON, J. P. S. (1964). Divergent thinking abilities in chronic schizophrenia. *J. clin. Psychol.* **20**, 433–435.

BANNISTER, D. (1960). Conceptual structure in thought-disordered schizophrenics. *J. ment. Sci.* **106**, 1230–1249.

BANNISTER, D. (1962). The nature and measurement of schizophrenic thought disorder. *J. ment. Sci.* **108**, 825–842.

BANNISTER, D. (1963). The genesis of schizophrenic thought disorder: a serial invalidation hypothesis. *Br. J. Psychiat.* **109**, 680–686.

BANNISTER, D. (1965). The genesis of schizophrenic thought disorder: retest of the serial invalidation hypothesis. *Br. J. Psychiat.* **111**, 377–382.

BANNISTER, D. & FRANSELLA, F. (1966). A grid test of schizophrenic thought disorder. *Br. J. soc. clin. Psychol.* **5**, 95–102.

BANNISTER, D. & FRANSELLA, F. (1967). *The Grid Test of Schizophrenic Thought Disorder: A Standard Clinical Test.* Barnstaple, Devon: Psychological Test Publications.

BANNISTER, D. & FRANSELLA, F. (1971). *Inquiring Man: A Psychology of the Person.* Harmondsworth: Penguin Books (in press).

BANNISTER, D. & MAIR, J. M. M. (1968). *The Evaluation of Personal Constructs.* London and New York: Academic Press.

BANNISTER, D. & SALMON, P. (1966). Schizophrenic thought disorder: specific or diffuse? *Br. J. med. Psychol.* **39**, 215–219.

FOULDS, G. A. *et al.* (1967). Cognitive disorder among the schizophrenics: I–II. *Br. J. Psychiat.* **113**, 1361–1374.

KELLY, G. A. (1955). *The Psychology of Personal Constructs*, vols. 1–2. New York: Norton.

KELLY, G. A. (1969). *Clinical Psychology and Personality: The Selected Papers of George Kelly*, edited by B. A. Maher. New York: Wiley.

MCPHERSON, F. M. (1969). Thought-process disorder, delusions of persecution and 'non-integration' in schizophrenia. *Br. J. med. Psychol.* **42**, 55–57.

MCPHERSON, F. M. *et al.* (1970). Flattening of affect and personal constructs. *Br. J. Psychiat.* **116**, 39–43.

PRESLY, A. S. (1969). 'Slowness' and performance on the grid test for thought disorder. *Br. J. soc. clin. Psychol.* **8**, 79–80.

ROMNEY, D. (1969). The validity of certain tests of overinclusion. *Br. J. Psychiat.* **115**, 591–592.

Manuscript received 8 *January* 1970

Revised manuscript received 18 *February* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 152–159
With 1 text-figure
Printed in Great Britain

Rate of Information Processing in a Schizophrenic and a Control Group: The Effect of Increasing Task Complexity

By P. D. SLADE

Department of Psychology, Institute of Psychiatry, University of London

In recent years two cognitive models of schizophrenic dysfunction have been elaborated which involve the basic proposition that schizophrenic patients process information at a slower rate than non-schizophrenic controls (McGhie, 1965; Yates, 1966). Both models predict (1) that schizophrenic patients will make decisions at a slower rate than controls at all levels of task complexity, and (2) that their relative slowness will increase as a function of increasing task complexity. The available evidence is consistent with the first prediction but ambiguous with respect to the second.

All previous studies relating to the above issues have utilized *discrete* reaction time measures. It was hypothesized that a more sensitive test of the predictions would be provided by a continuous performance task involving responding over a longer time period. A study is reported in which a group of chronic schizophrenic patients was compared with a group of non-schizophrenic control patients on a series of four card-sorting tasks in which the 'information to be processed' was systematically varied. In general, the findings were consistent with expectations.

In recent years two cognitive models of schizophrenic dysfunction have been elaborated which involve the basic proposition that schizophrenic patients process information at a slower rate than non-schizophrenic controls. The names primarily associated with these two models are McGhie (1965) and Yates (1966). The two formulations differ with respect to the stage at which they place the primary defect. McGhie suggests that it is due to an inability of the schizophrenic patient to screen out task-irrelevant information which has to be processed along with task-relevant information, while Yates suggests that it is due to the abnormally slow rate at which the central channel can deal with task-*relevant* information. This difference between the two formulations, however, is not important for present purposes.

In order to test the proposition that schizophrenic patients process information at a slower rate than non-schizophrenics, it is necessary to have some means of quantifying the amount of information provided by any stimulus or that involved in any type of task performance. The communication theory equations developed by Shannon in his now famous 1949 paper were selected for this purpose. Shannon suggested that in the simple situation in which a number of equiprobable stimuli are possible, the amount of information provided by the occurrence of any one of them can be described by the base 2 logarithm of the number of stimuli. Similarly, in the situation in which a series of equiprobable stimuli occur, the total 'information' provided by the series can be described by the sum of the 'information' values of each of the stimuli in the series. When this definition is used, the unit of information is termed a 'bit'.

Having briefly indicated the basic approach to the measurement of task information, it is necessary to consider the method used for estimating information processing rates. The concept of channel capacity is important in this respect. It was assumed (with some justification, e.g. Hick, 1952) that individuals process information at a constant rate, and that this rate (or estimate of channel capacity) can be estimated from a knowledge of two variables, namely (1) the amount of 'information' in the task, and (2) the time required for the individual to carry out the task. The relationship between these two variables is shown in Fig. 1 for two hypothetical decision channels, which are capable of handling information at different rates. It can be seen from Fig. 1 that the smaller capacity channel exhibits a steeper time-complexity slope. When viewed in these terms the proposition that schizophrenic patients process information at a slower rate than non-schizophrenics gives rise to at least two predictions, namely (1) that schizophrenic patients make decisions at a slower rate than non-schizophrenics at *all* levels of task complexity, and (2) that their relative slowness will increase as a function of increasing task complexity.

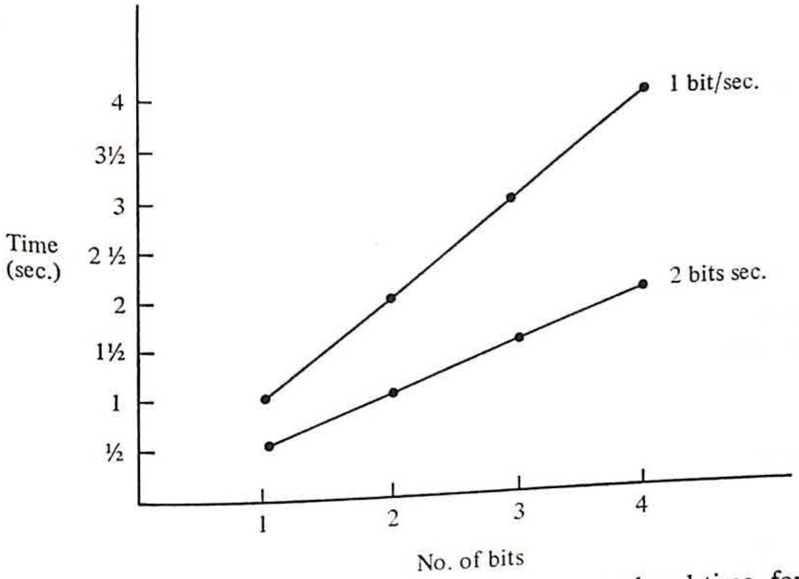


Fig. 1. The relationship between number of bits to be processed and time, for two hypothetical decision channels.

Several studies have been carried out which relate to these predictions (e.g. King, 1954; Venables, 1958; Harwood & Naylor, 1963; Karras, 1967; Court & Garwoli, 1968). The findings of all of these studies have been consistent with the first prediction, while discrepant findings have emerged with respect to the prediction of differential time-complexity curves. However, all of these studies have utilized *discrete* reaction-time tasks. It was hypothesized that a more sensitive test of the predictions might be provided by a continuous performance task, than by a discrete RT task, for two reasons: (1) It has been reported by some investigators that the discrete RT task fails to show up differences in decision time with different levels of task complexity (e.g. Mowbray & Rhoades, 1959; Leonard, 1959). (2) If the 'irrelevant-distraction' model of McGhie is valid, it is just as likely that distraction occurs on a *non-continuous* as on a continuous basis. If this is so, a discrete RT task involving responses of a fraction of a second may fail to show the deficit.

PRESENT STUDY

Materials and apparatus. The technique adopted for measuring 'information processing rates' was based on a modification of the card-sorting task used by Crossman (1953). Four packs of cards were prepared, with 32 cards in each pack. The first pack comprised equal numbers of two categories (i.e. 16 cards/category); the second pack comprised equal numbers of four categories (i.e. 8 cards/category); the third pack comprised equal numbers of eight categories (i.e. 4 cards/category) and the fourth pack comprised equal numbers of 16 categories (i.e. 2 cards/category). According to Shannon's equations, when each pack is thoroughly shuffled, the total information in each pack is as follows: Pack 1, 32 bits; Pack 2, 64 bits; Pack 3, 96 bits; Pack 4, 128 bits.

The problem of finding appropriate visual stimuli for the various categories posed some difficulties in the light of Crossman's findings concerning stimulus discriminability (Crossman, 1955). It was finally decided to use what are generally considered to be the primary visual dimensions: namely shape, colour and figure-ground. Each card was 3 in. square and the four shapes used were equated in terms of surface area, each shape yielding one square inch in area. No similar control, however, was provided for colour saturation.

The sorting apparatus consisted of a raised board $15\frac{1}{4}$ in. square, on which 16 boxes, each $3\frac{1}{2}$ in. square, were constructed. Each box was surrounded by a wall $\frac{1}{4}$ in. wide and $\frac{1}{2}$ in. in depth, which served to separate it from its neighbour. The apparatus, which was painted black, rested on a table approximately 2 ft. 6 in. high.

Procedure. Following Crossman, it was decided to obtain independent measures of 'simple movement time' and 'total sorting time' so that a measure of 'decision time' could be calculated by subtraction. The procedure therefore involved two stages.

Stage 1 (measurement of simple movement time). Subjects were required to sort the packs of 32 cards into 2, 4, 8 and 16 boxes as quickly as possible, with instructions to ignore the design on each card. White marker cards were used to indicate the number and location of boxes into which the subject was required to sort on each trial. Following a demonstration trial, subjects were given one practice trial on each of the four conditions followed by one test trial.

Stage 2 (measurement of total sorting time). Subjects were then required to sort each of the four packs of cards into the appropriate categories, with the aid of patterned marker cards as quickly as possible. Two test trials were given for each pack, an ABBA design being employed.

Experimental group. The experimental group comprised a pool of 16 chronic schizophrenic patients, all of whom had, on their first admission to hospital, been diagnosed as suffering from a schizophrenic illness, this diagnosis having been confirmed on all subsequent examinations. Seven of the patients had been diagnosed as predominantly paranoid, while nine had been diagnosed as predominantly non-paranoid. With one exception, all patients were receiving maintenance dosages of medication, mainly phenothiazines, the highest recorded dosage being chlorpromazine 100 mg. TDS. Twelve of these patients were matched as closely as possible, on an individual basis, for age and Progressive Matrices IQ with the 12 control patients (see Table 1).

Control group. The control group comprised 12 non-psychotic, non-organic psychiatric patients, who, with one exception, fell within the psychiatric category of psychoneurotic. The most frequent descriptive labels were reactive depression and chronic anxiety state. With one exception, all patients were receiving medication, mainly antidepressants.

RESULTS

Simple movement times. Mean movement times for the schizophrenic and control groups are shown in the upper half of Table 2. The means and standard deviations were found to be related, necessitating a logarithmic transformation of individual

Table 1. *General characteristics of the two matched groups*

	Chronic schizophrenics ($n=12$)	Control patients ($n=12$)	*t* test
Age			n.s.
Mean	40·75	42·08	
S.D.	8·12	10·34	
Range	25–56	22–60	
P.M.I.Q.			n.s.
Mean	107·17	107·66	
S.D.	11·58	12·60	
Range	88–126	92–126	
Education (yr.)			n.s.
Mean	10·17	9·83	
S.D.	2·02	2·42	
Range	9–16	8–13	

times. The transformed times were then subjected to an analysis of variance. The results of this procedure indicated a significant difference between groups ($F=8{\cdot}99$; $P<0{\cdot}01$), no significant differences between conditions ($F=2{\cdot}06$; $P>0{\cdot}10$) and no significant groups × conditions interaction ($F=0{\cdot}673$; $P>0{\cdot}25$). With respect to the simple movement time procedure, therefore, the chronic schizophrenics were significantly slower than the controls, while the factor of complexity was not found to be important (i.e. with the materials and apparatus used in the present study, subjects were generally able to place cards randomly into 16 boxes as quickly as they could into two boxes).

Decision times. Since, in the analysis of simple movement times, no significant differences were found between 'complexity conditions', it was decided to take each patient's average movement time over the four conditions as the best estimate of his movement time. Each patient's decision time was then computed by subtracting this estimate of movement time from his total sorting time for each condition (an average of the two trials that were given). The mean decision times for the two groups, calculated in this way, are shown in the lower half of Table 2.

Table 2. *Mean sorting times for the two matched groups*

	Mean simple movement times (sec.)			
No. of boxes	2	4	8	16
Chronic schizophrenics ($n=12$)	32·50	31·92	33·92	33
Control patients ($n=12$)	21·83	22·25	22·60	23·25

	Mean decision times (sec.)			
No. of categories	2	4	8	16
Chronic schizophrenics ($n=12$)	16	23·17	37·97	60
Control patients ($n=12$)	13·65	20·31	29·12	48·94

Analysis in terms of 'information processing rates'. A measure of 'processing rate' (a bits/sec. index) was derived for every patient, for each of the four complexity conditions, by means of the formula:

$$\text{No. of bits/sec.} = \frac{\text{No. of bits to be processed}}{\text{Decision time in sec.}}$$

The bits/sec. indices were then subjected to an analysis of variance to test for 'group' and 'complexity conditions' differences. The results of this analysis are shown in Table 3. It can be seen from Table 3 that the analysis of variance indicated a significant difference between the two groups. This group difference was found to be significant for all conditions, thus confirming the prediction that chronic schizophrenics make decisions at a slower rate than controls at all levels of task complexity. Also, since the bits/sec. index is a ratio index, this means in effect that the curve relating decision time to task complexity is significantly steeper for the schizophrenic patients than the controls, thus confirming the second prediction.

Table 3. *Analysis of variance of bits/sec. indices*

Source	D.F.	S.S.	M.S.	V.R.	Significance
Groups	1	5·0325	5·0325	23·45	$<0{\cdot}005$
Conditions	3	10·4115	3·4705	21·94	$<0{\cdot}005$
G × C	3	0·5638	0·1879	1·18	n.s.
Subjects within groups	22	4·7211	0·2146	1·356	n.s.
Residual	66	10·4404	0·1582		
Total	95	31·1693			

The analysis of variance also indicated significant differences between complexity conditions. A Tukey multiple range test showed that the significant differences lay between the two extreme and the two intermediate complexity levels. That is, patients achieved significantly higher rates of information processing (as here defined) with the 4- and 8-category conditions. Furthermore, the groups × conditions interaction was not significant, suggesting that both patient groups exhibited a similar pattern of response with respect to the four levels of complexity. To put it another way, the decision times of both groups differ significantly from linearity when plotted against a logarithmic function of task complexity.

Linear regression equation analysis. In view of the finding that the decision-time data diverged significantly from a logarithmic function of task complexity, it was decided to see what mathematical function would provide a better fit for the group curves. From a visual inspection of the data plotted against an arithmetic function of number of categories, it seemed that decision times fit an arithmetic function better than a logarithmic function. Straight line regression equations were therefore fitted to the decision-time data of all patients, by means of the formula $Y = a + bX$, where Y equals decision time and X equals number of categories. The mean regression equations for the two groups were:

Chronic schizophrenics ($n = 12$): $Y = 10{\cdot}803 + 3{\cdot}13X$.
Control patients ($n = 12$): $Y = 9{\cdot}487 + 2{\cdot}47X$.

Correlated *t* tests were then used to test for group differences on the *a* and *b* regression coefficients separately. The results indicated that the two groups did not differ significantly on the *a* coefficient, the *constant* component of decision time ($t = 0{\cdot}88$; $P > 0{\cdot}25$), while they did differ significantly on the *b* coefficient, the *ratio* component relating decision time to task complexity ($t = 7{\cdot}22$; $P < 0{\cdot}01$). The latter significant difference again provides support for the two predictions.

Chronic paranoid v. chronic non-paranoid dichotomy. The original pool of 16 chronic schizophrenics were split into a predominantly paranoid ($n = 7$) and a predominantly non-paranoid group ($n = 9$) and compared, along with the group of control patients ($n = 12$), on some of the measures.

An analysis of variance of the simple movement time scores revealed significant differences between groups at the 5 per cent level; but when Student's *t* test was carried out, it was found that the significant difference lay between both schizophrenic groups and the control group, while there was no significant difference between the paranoid and non-paranoid subgroups.

An analysis of the regression equation coefficients, by means of *t* tests, indicated no significant differences between groups on the *a* coefficients, while only the non-paranoid schizophrenic group differed significantly from the controls in terms of the *b* coefficients ($t = 2{\cdot}55$; $P < 0{\cdot}02$).

DISCUSSION

The finding of significantly slower movement times in the chronic schizophrenic group is consistent with the findings of McGhie *et al.* (1965*b*) and with those of other studies indicating motor slowness in this type of patient: as such, it does not require further discussion. In contrast, the findings concerning decision times need rather more elaboration.

The first measure examined, the bits/sec. index, is in effect a ratio index. It represents the ratio of task complexity to decision time, where task complexity is defined as the logarithmic function of degree of choice (i.e. number of categories). The *larger* the index the *faster* the rate of information processing. The major hypothesis tested was that the group of schizophrenic patients would have a significantly smaller average index than the control group. The analysis of variance confirmed this hypothesis and the two predictions derived from it. One aspect of the communication theory analysis, however, was unexpected: namely the finding of significant differences between conditions (degrees of complexity). Communication theory suggests that, while individuals may vary in terms of the magnitude of the index, for any one individual it will tend to be fairly constant over a wide range of task complexity. The analysis of variance revealed that this was not true in relation to the present results. Two possible explanations of this finding suggest themselves. Firstly, that the criterion of equal stimulus discriminability was not achieved in constructing the sorting tasks used (although, following Crossman, 1955, an attempt was made to control for stimulus size), and that the discrepancies from a logarithmic function are due to this fact. Secondly, that with the present tasks, the present range of task-complexity, and the present groups of patients, a logarithmic function is not the most appropriate one for determining the amount of information in the task. If the first explanation is the correct one, and it seems very probable that it is, it could be argued that the differences between the two groups in terms of time-complexity curves involve a confound between complexity differences and differences in discriminability between categories in the four conditions. Or, to put it another way, that the differences between the slopes of the two groups are due to perceptual rather than complexity differences. In response to this argument, it was found that there were no significant differences in terms of the bits/sec. indices between the 2- and 16-category conditions and between the 4- and 8-category conditions. That is, there was no evidence to suggest that the 2- and 16-category conditions do not fit the same logarithmic curve and therefore differ in terms of discriminability; nor, similarly, that the 4- and 8-category conditions do not fit the same logarithmic curve and differ in terms of discriminability. Furthermore, it was found that the differences between groups (in terms of the bits/sec. indices) existed at *all* four levels (i.e. with all four conditions), and not simply between levels which may be said to differ in terms of discriminability. A further contrary argument stems from the pattern shown by the two groups with respect to the four complexity levels. Both groups produced faster rates of 'information processing' with the two intermediate levels when compared with the two extreme levels. If the two intermediate and two extreme levels differ in discriminability and the differences between the schizophrenic and control groups were due to this fact, one would expect a different pattern of response for the two groups rather than a similar one. It seems clear, therefore, that while differences in discriminability may be involved in the tasks used, these differences are not responsible for the differences between the groups.

Straight line regression equations were also fitted to the decision-time data as a result of a visual inspection, which indicated that an arithmetic function fitted the present data better than a logarithmic one. The *b* coefficient, or ratio component of decision time, seems fairly easy to interpret. It represents the direct relationship between task complexity and decision time. Therefore the finding of significantly larger *b* coefficients for the schizophrenic patients again supports the major hypothesis. The *a* coefficients, or constant component of decision time, are more difficult to understand.

The fact that for all patients the *a* coefficients were positive and greater than zero contraindicates a simple interpretation in terms of error due to an inexact fit. Similarly, the possibility that it represents some aspect of motor-response time can be excluded on the grounds that simple movement times had already been subtracted from total sorting times in order to derive the decision-time data. A possible interpretation is that it represents 'neural transmission time' or some constant lag involved in continuous performance tasks.

With respect to the analysis of subgroup differences, the finding that only the non-paranoids differed significantly from the controls on the *b* coefficients suggests that this subgroup is more affected by the complexity variable. This finding parallels that of McGhie *et al.* (1965*a*) that their hebephrenic (or non-paranoid) patients were the ones most prone to the effects of experimental distraction.

CONCLUSIONS

The results of the present study have been analysed in two different ways, both of which point to the importance of a complexity factor in determining speed of response in chronic schizophrenics. The essential conclusion would seem to be that, despite the failure to provide an adequate control for stimulus discriminability, evidence has been provided for the proposition that chronic schizophrenic patients process information at a slower rate than controls. Support was given to both the predictions derived from this proposition, namely (1) that chronic schizophrenic patients make decisions at a slower rate than controls at all levels of task complexity, and (2) that they make decisions at an increasingly slower rate, compared with controls, as task complexity increases.

From a clinical point of view, this would seem to suggest that chronic schizophrenic patients may be at their best with relatively simple tasks and exhibit increasing speed deficit as task complexity increases.

The study reported here was part of a larger investigation conducted in partial fulfilment of the degree of Master of Philosophy of the University of London. This paper is a slightly revised version of a paper presented at the London Conference of the British Psychological Society in December 1969.

REFERENCES

Court, J. H. & Garwoli, E. (1968). Schizophrenic performance on a reaction-time task with increasing levels of complexity. *Br. J. soc. clin. Psychol.* **7**, 216–223.

Crossman, E. R. F. W. (1953). Entropy and choice time: the effect of frequency unbalance on choice-response. *Q. J. exp. Psychol.* **5**, 41–51.

Crossman, E. R. F. W. (1955). The measurement of discriminability. *Q. J. exp. Psychol.* **7**, 176–195.

Harwood, E. & Naylor, G. F. K. (1963). Nature and extent of basic cognitive deterioration in a sample of institutionalized mental patients. *Aust. J. Psychol.* **15**, 29–36.

Hick, W. E. (1952). On the rate of gain of information. *Q. J. exp. Psychol.* **4**, 11–26.

Karras, A. (1967). The effect of stimulus response complexity on the reaction time of schizophrenics. *Psychon. Sci.* **7**, 75–76.

King, H. E. (1954). *Psychomotor Aspects of Mental Disease.* Cambridge, Mass.: Harvard University Press.

Leonard, J. A. (1959). Tactual choice reactions. *Q. J. exp. Psychol.* **11**, 76–83.

McGhie, A. (1965). Psychological studies of schizophrenia. In T. Freeman, J. L. Cameron & A. McGhie (eds), *Studies on Psychoses.* London: Tavistock Publications.

McGhie, A., Chapman, J. & Lawson, J. S. (1965*a*). The effect of distraction on schizophrenic performance: 1. Perception and immediate memory. *Br. J. Psychiat.* **111**, 383–390.

McGhie, A., Chapman, J. & Lawson, J. S. (1965*b*). The effect of distraction on schizophrenic performance: 2. Psychomotor ability. *Br. J. Psychiat.* **111**, 391–398.

Mowbray, G. H. & Rhoades, M. V. (1959). On the reduction of choice reaction times with practice. *Q. J. exp. Psychol.* **11**, 16–23.

Shannon, C. E. & Weaver, W. (1949). *The Mathematical Theory of Communication.* Urbana, Ill.: University of Illinois Press.

Venables, P. H. (1958). Stimulus complexity as a determinant of the reaction time of schizophrenics. *Can. J. Psychol.* **12**, 187–190.

Yates, A. J. (1966). Data processing levels and thought disorder in schizophrenia. *Aust. J. Psychol.* **18**, 103–117.

Manuscript received 12 February 1970

Revised manuscript received 2 July 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 160–171
With 1 text-figure
Printed in Great Britain

'Psychoticism' in Drug-Users

BY JOHN D. TEASDALE
Department of Psychology, Institute of Psychiatry, University of London, and Drug Dependence Clinical Research and Treatment Unit, Maudsley Hospital and Bethlem Royal Hospital

R. TAYLOR SEGRAVES*
Department of Psychology, Institute of Psychiatry, University of London

AND JIM ZACUNE
Addiction Research Unit, Institute of Psychiatry, University of London

Scores on the PEN Inventory were obtained from four groups of drug-users. They were compared with scores from a group of apprentices who had completed the Inventory non-anonymously rather than the normative data, as the latter were obtained anonymously, and it was found that this factor significantly elevated P ('psychoticism') scores. All four groups were significantly higher on P than the comparison groups. This was not simply the result of endorsement of items directly related to drug-taking, nor to the correlation of P and N scores. Item analysis showed that 11 of the 20 items significantly differentiated drug-using and control groups, but item content was more homogeneous in the items which did not differentiate the groups. Hypotheses to account for the high P scores of drug-users are suggested.

Three of the drug-using groups were significantly high on neuroticism, the fourth approaching significance in that direction. Two of the drug-using groups were significantly low on extraversion; the other two did not differ from a control population.

Significant differences were obtained on P, E and N scores between the four drug-using groups.

In a number of papers, Eysenck & Eysenck (1968*a*, *b*, 1969) have described the development and use of a set of questionnaire items for the measurement of the P ('psychoticism') dimension of personality. The P dimension is conceived of as orthogonal to the previously established dimensions of E (extraversion) and N (neuroticism). It is hoped by its use to obtain a more comprehensive picture of an individual's position on factorially established personality dimensions than could be provided by E and N scores on their own.

That the P dimension might be of some importance in describing the personality of drug-users was suggested by the fact that the item 'Would you take drugs which may have strange or dangerous effects?' loads positively on the P factor in males, and is included as one of the 20 items used to measure P in the PEN Inventory (Eysenck & Eysenck, 1968*b*). It was accordingly predicted that drug-users would have elevated P scores.

---

* Present address: Medical School, Vanderbilt University, Nashville, Tennessee.

METHOD

The PEN Inventory consists of the 20 P, 20 E and 20 N items listed in Eysenck & Eysenck (1968*b*) together with the 18 items of the Lie scale of the Eysenck Personality Inventory, Forms A and B combined (Eysenck & Eysenck, 1964). The subject completing the Inventory is required to fill in his name, age and sex.

Scores were available on this inventory from the following four groups of male drug-users.

Provincial heroin-users (*group PH*)

This group consisted of 21 heroin-users living in an English provincial town. They were all those males of a larger group of heroin-users intensively studied for whom completed PEN Inventories were available. The characteristics and method of contacting this group have been fully described elsewhere (Kosviner *et al.*, 1968). Briefly, the group included both those using heroin regularly and those using it irregularly; most of the sample were living in the community (i.e. not hospitalized) at the time of the study. Their mean age was 20 years 8 months (S.D. 3·5 years).

In-patient injectors (*group IPI*)

The 22 males in this group had been admitted as voluntary patients to the closed ward of the Drug Dependence Clinical Research and Treatment Unit, Bethlem Royal Hospital, for the treatment of drug dependence. All patients had administered drugs by the intravenous route; at the time of admission the main drug of dependence was heroin for 17, and methedrine for five. This group consisted of all heroin- and methedrine-users admitted on whom scores on the PEN Inventory were available.

Although the two subgroups actually presented the injection of different drugs as their problem, they were placed together in the in-patient injector group, as their means on the P score were practically identical (heroin, 4·47; methedrine, 4·40) and also the distinction in terms of main drug of dependence was not very sharp, i.e. some of the heroin subgroup had used methedrine, and some of the methedrine subgroup had used heroin.

Their mean age was 20 years 8 months (S.D. 2·5 years).

In-patient oral (*group IPO*)

This group consisted of 11 males admitted as voluntary patients to the open ward of the Drug Dependence Clinical Research and Treatment Unit, Bethlem Royal Hospital, for the treatment of dependence on orally administered stimulants. The main drug of dependence was Drinamyl in nine of the 11 cases, and amphetamines in the other two.

Their mean age was 21 years 4 months (S.D. 4·4 years). These 11 were all the patients of this type on whom PEN scores were available.

Experimental (*group EX*)

This group of 10 subjects was obtained as part of a larger study (Segraves, 1969) involving the measurement of urinary corticosteroid levels. To make this measurement accurately it was necessary to ensure that the subjects were not using drugs. Accordingly, in a questionnaire that all experimental subjects completed, a question asking whether the subject had taken any LSD, marijuana or amphetamines in the preceding three weeks was included. Of the total of 100 subjects completing the questionnaire and the PEN Inventory, 10 claimed to have taken one of the above drugs in the preceding three weeks. These 10 constitute the experimental group. That there was some validity in their claim that they took such drugs is suggested by the fact that two of these 10 were known later to have been arrested by the police and charged with illegal possession of drugs.

These subjects were students attending a technical college in north London. The mean age of the experimental group was 19 years 10 months (S.D. 0·9 years).

Comparison groups

Eysenck & Eysenck (1969) report that scores on P vary appreciably with age and social class. Thus it is important to get a control group of non-drug-users of similar age and social class to each group of drug-users, against which to compare them, rather than comparing them with the mean for the population as a whole.

For group EX, the 90 non-drug-using subjects were used as the comparison group as they came from the same population, defined by attendance at the same technical college, as the 10 drug-users.

For the other three groups, it was intended to obtain comparison groups by selecting for each drug-user a subject of the same age and social class from the normative data reported by Eysenck & Eysenck (1969). (These data were very kindly made available by Dr S. B. G. Eysenck.) When this was in fact done, it appeared that these comparison groups had means somewhat higher than that for the 90 subjects forming the EX comparison group. It was suspected that this was related to the fact that subjects completing the inventory in the Eysenck & Eysenck study were not required to put their names on the questionnaires, whereas those in the EX comparison group were. This hypothesis was tested by comparing the scores of two groups of subjects from the same population (Ford apprentices), the groups differing only in that one group completed the questionnaire anonymously while the other filled in their names. (These data were also very kindly made available by Dr S. B. G. Eysenck.) The mean P score of the group completing the questionnaire anonymously was 3·08 ($n=75$) and of the other group 2·13 ($n=210$). Both distributions were highly skewed, the modal score being 0 for both groups. Accordingly, a χ^2 test was used to compare the two distributions, dichotomizing the distributions between a score of 4 and a score of 5. The resulting χ^2 was 7·31, significant beyond the $P=0{\cdot}01$ level. The hypothesis was thus confirmed, and it was considered inappropriate to employ the comparison groups which had already been assembled as they, unlike the drug-using groups, had completed the questionnaire anonymously.

Lacking a better comparison group of subjects who had completed the PEN Inventory with their names on it, it was decided to use the group of 210 Ford apprentices. The mean age of this group was 17 years 1 month (S.D. 0·8 years), and they were all social class 5, in the classification used by Eysenck & Eysenck (1969). This group was both younger and of lower average social class than the drug-using groups. From the data of Eysenck & Eysenck (1969), it would be predicted that the effects of these age and class differences would be to make the P score of the comparison group higher than that of the drug-users, in as much as these variables are affecting the P score. Thus if the drug-users are found to have higher P scores than the comparison group, it cannot be argued that this results simply from the differences in age and social class between them. Hence the use of this comparison group seems justifiable.

That the comparison group is relatively high on P compared with other normal groups is shown by the fact that the mean for the group of apprentices who filled in the questionnaire anonymously, and who are thus comparable with the data reported by Eysenck & Eysenck (1969), lies above those of 16 of the 21 occupational groups reported in that paper. It thus seems unlikely, if the drug-users score significantly higher on P than the comparison group, that this results simply from using a group relatively low on P as a control group.

RESULTS

Comparison of drug-using and control groups

The distributions of P scores for the two comparison groups and the four drug-using groups are shown in Fig. 1. The means are shown in Table 1; because of the non-normality of the distributions standard deviations are not given. Fisher's exact probability test was used to compare the distributions of P in the drug-using groups with the control groups. In all cases, the distributions were dichotomized between a P score of 1 and a P score of 2; this was the closest approximation to a median split.

It can be seen that all the drug-using groups are significantly higher on P than their comparison groups. However, it is possible that this difference might simply result from an increased rate of endorsement of the P item 'Would you take drugs which may have strange or dangerous effects?' in the drug-using groups. If this

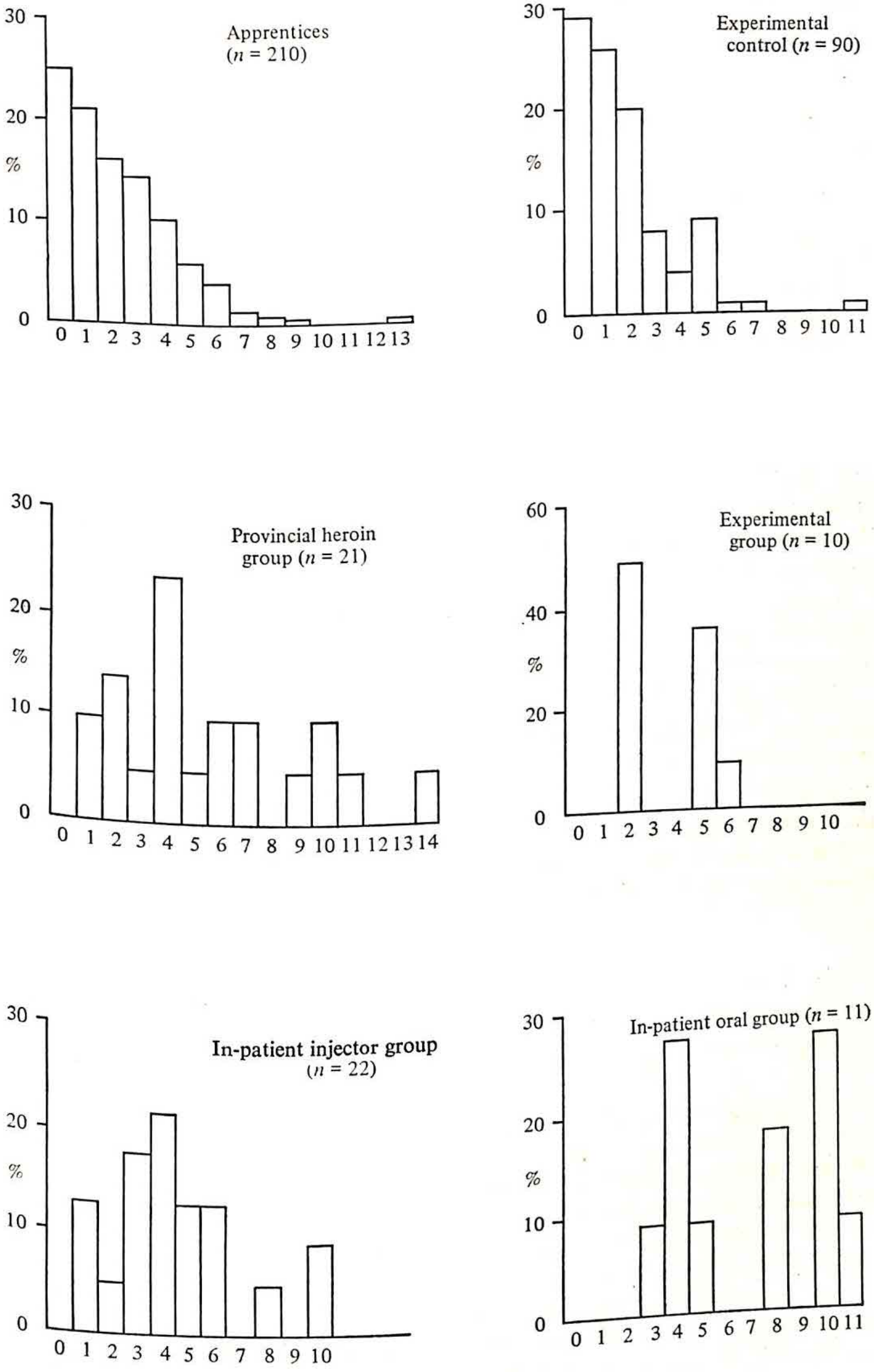


Fig. 1. Distribution of P scores for drug-using and control groups.

Table 1. *Mean P scores and results of Fisher's exact probability tests comparing drug-using groups with control groups*

Group	*n*	Mean	Comparison group	Fisher's exact probability*	Point of dichotomization
(*a*) Apprentices	210	2·13	—		
(*b*) Experimental comparison	90	1·82	—		
Experimental	10	3·60	*b*	$P=0{\cdot}00059$	1/2
Provincial heroin	21	5·48	*a*	$P=0{\cdot}00045$	1/2
In-patient injectors	22	4·46	*a*	$P=0{\cdot}0024$	1/2
In-patient oral	11	6·91	*a*	$P=0{\cdot}0011$	1/2

* One-tailed test.

were the case, then the elevated P score in the drug-using groups would not be of much significance in describing the personality of drug-users. Accordingly, the P scores of the six groups were compared again after eliminating the effect of this item from the total P score. The results are shown in Table 2.

It can be seen that there are still significant differences between three of the drug-using groups and their respective comparison groups, the fourth (EX) being only marginally significantly different.

To gain a picture of what items contributed to the higher P score in the drug-using groups, an item analysis was performed. In order to get sufficient numbers to make this feasible, it was necessary to combine all four drug-using groups together. 2×2 tables were then constructed for the frequency of 'yes' and 'no' responses in the combined drug-using group and in the apprentice control group.

Table 3 shows the items which differ significantly ($P<0{\cdot}05$) in endorsement frequency for the two groups, and those which did not. In all cases where there is a

Table 2. *Mean P scores and results of Fisher's exact probability tests comparing drug-using groups with control groups, omitting the effect of the item 'Would you take drugs which may have strange or dangerous effects?' from the total P score*

Group	*n*	Mean	Comparison group	Fisher's exact probability*	Point of dichotomization
(*a*) Apprentices	210	2·06	—		
(*b*) Experimental comparison	90	1·68	—		
Experimental	10	2·60	*b*	$P=0{\cdot}031$	0/1
Provincial heroin	21	4·48	*a*	$P=0{\cdot}018$	1/2
In-patient injectors	22	3·68	*a*	$P=0{\cdot}0013$	1/2
In-patient oral	11	6·36	*a*	$P=0{\cdot}00066$	1/2

* One-tailed test.

significant difference between the two groups, it results from a higher frequency of endorsement in the P direction by the drug-using group.

Eleven of the 20 P items differ significantly in their frequency of endorsement between the two groups. It could be argued that five of these items (nos. 12, 14, 15, 16 and 18 in Table 3) were endorsed in the P direction by the drug-users simply as a result of their being drug-users, rather than as a result of their being high on P before starting to use drugs. It is thus important to check whether differences between the drug-using groups and the control group still exist when these items are eliminated from the calculation of the P score. The results of this procedure are shown in Table 4. Groups IPO and IPI are significantly different from their control groups, dichotomizing as close as possible to a median split, employing a one-tailed test. The EX group just fails to reach the 5 per cent level of significance when the same procedure is employed. The PH group, while not significantly different from its control group on a median split, is highly significantly different from it when the groups are dichotomized between P scores of 5 and 6, suggesting that

Table 3. *Item analysis of P scale in relation to differences between drug-using and control group*

(*a*) Items not differing significantly in frequency of endorsement between drug-using and control group:

1. Are there several people who keep trying to avoid you? ($P < 0·1$)
2. Is there someone who is responsible for most of your troubles?
3. Do people generally seem to take offence easily?
4. Do you have enemies who wish to harm you?
5. Do your friendships break up easily without it being your fault? ($P < 0·1$)
6. Do people mean to say and do things to annoy you?
7. Would you have been more successful if people had not put difficulties in your way?
8. Would it upset you a lot to see a child or animal suffer?
9. When you are in a crowded place like a bus, do you worry about dangers of infection?

(*b*) Items differing significantly in frequency of endorsement between drug-using and control group:

Item	Chi-square	*P*	% drug-using group scoring in P direction
10. Do most things taste the same to you?	9·02	<0·01	17
11. Do you enjoy hurting people you love?	12·53	<0·001	19
12. Are you generally in good health?	35·82	<0·001	20
13. Was your mother a good woman?	12·23	<0·001	9
14. Have you had more trouble than most?	41·12	<0·001	50
15. Do you worry a lot about catching diseases?	6·66	<0·01	28
16. Have you had an awful lot of bad luck?	25·01	<0·001	45
17. Do you let your dreams warn or guide you?	13·69	<0·001	28
18. Would you take drugs which may have strange or dangerous effects?	136·99	<0·001	83
19. Did you love your mother?	16·96	<0·001	17
20. Was your father a good man?	18·85	<0·001	16

Table 4. *Mean P scores and results of Fisher's exact probability tests comparing drug-using groups with control groups omitting items 12, 14, 15, 16 and 18. (See Table 3.)*

Group	*n*	Mean	Comparison group	Fisher's exact probability*	Point of dichotomization
(*a*) Apprentices	210	1·61	—		
(*b*) Experimental control	90	1·29	—		
Experimental	10	2·00	*b*	$P = 0·059$	0/1
Provincial heroin	21	2·86	*a*	$P = 0·00067$	5/6
In-patient injectors	22	2·32	*a*	$P = 0·028$	1/2
In-patient oral	11	4·36	*a*	$P = 0·016$	1/2

* One-tailed test.

differences at the upper end of the P distribution are more marked than those at the lower end.

The data give some support to the suggestion that drug-using groups are still higher on P than their control groups even when the effects of five of the 11 items which significantly differentiated drug-users and non-users have been eliminated.

Scores on E and N

Scores on E, N and L for the drug-using and control groups are shown in Table 5. It can be seen that the two control groups that have been employed in this study are quite similar on the variables of extraversion and neuroticism to the normative group reported by Eysenck & Eysenck (1969). The group of apprentices does not differ significantly from this normative group on neuroticism and while it is more extraverted to a statistically significant level ($t = 3·64$; $P < 0·001$), this difference is very slight, amounting to approximately $\frac{1}{4}$ S.D. of the normative group.

The EX control group does not differ significantly from the normative group on extraversion, but does on neuroticism ($t = 2·91$; $P < 0·01$), though here again the difference is only slight, amounting to less than $\frac{1}{3}$ S.D. of the normative group.

The two control groups employed thus seem reasonably representative of the

Table 5. *Scores on E, N and L from the PEN Inventory*

Group	E		N		L	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Normative group* ($n = 843$)	13·46	3·83	8·34	4·30		
Apprentices ($n = 210$)	14·41	3·27	8·86	3·84	3·14	2·07
Experimental control ($n = 90$)	13·11	4·41	9·68	4·14	3·20	2·41
Experimental ($n = 10$)	15·00	3·53	12·20	4·13	2·60	2·59
Provincial heroin ($n = 21$)	10·00	4·89	15·00	3·44	2·81	2·73
In-patient injectors ($n = 22$)	12·82	4·34	11·64	4·76	4·00	2·69
In-patient oral ($n = 11$)	10·09	4·32	16·00	2·83	3·82	1·60

* This is the 10–29 age-group reported in Eysenck & Eysenck (1969).

younger section of the general population on the variables of extraversion and neuroticism.

The EX group is not significantly different from the EX control group on either extraversion ($t=1{\cdot}29$, n.s.) or neuroticism ($t=1{\cdot}83$; $P<0{\cdot}1$), although it approaches being significantly higher on the latter.

The PH group is significantly low on extraversion, both when compared with the group of apprentices ($t=4{\cdot}04$; $P<0{\cdot}001$) and the normative group ($t=3{\cdot}22$; $P<0{\cdot}002$). Further, the differences are of the order of 1 S.D. of the control groups. On neuroticism, the PH group is significantly higher than either the group of apprentices ($t=7{\cdot}71$; $P<0{\cdot}001$) or the normative group ($t=8{\cdot}70$; $P<0{\cdot}001$). Here the difference from the control groups is approximately 1½ S.D. of the control groups.

The IPI group does not differ significantly on extraversion from either the apprentices or the normative group. On neuroticism, however, it is significantly higher than both the apprentices ($t=2{\cdot}65$; $P<0{\cdot}01$) and the normative group ($t=3{\cdot}22$; $P<0{\cdot}002$). These differences are of the order of ¾ S.D. of the control groups.

Group IPO is significantly lower on extraversion than both the apprentices ($t=3{\cdot}27$; $P<0{\cdot}002$) and the normative group ($t=2{\cdot}57$; $P<0{\cdot}002$), the difference approximating 1 S.D. of the control groups. On neuroticism, these patients are significantly higher than the apprentices ($t=7{\cdot}99$; $P<0{\cdot}001$) and the normative group ($t=8{\cdot}85$; $P<0{\cdot}001$), to the extent of at least 1½ S.D. of the control groups.

Correction of P scores for correlation with N

Eysenck & Eysenck (1968*b*) report a product-moment correlation of 0·45 between N and P for their normal male sample. In the present study this correlation in the two control groups was 0·38 for the apprentices ($P<0{\cdot}001$) and 0·35 ($P<0{\cdot}001$) for the EX control group. It is thus possible that the elevated P scores in the drug-using groups result simply from these groups having elevated N scores.

To examine this possibility linear regression equations to calculate the values of P predicted from the regression on N in the normal population were derived for the two control groups separately. By inserting their N scores in these equations, predicted values of P for each drug-user were calculated, using the equation derived from the appropriate control group. These predicted values of P were compared with those actually obtained for each group of drug-users by means of correlated t tests. It was found that the P values actually obtained were significantly higher (beyond the $P=0{\cdot}02$ level) than those predicted from the N scores for all four groups. The elevated P scores of the drug-users could thus not be accounted for solely in terms of their elevated N scores.

A final stage was to see whether the elevated P scores of the drug-users could be explained by the combined effects of elevated N scores, and of endorsement of items directly related to the results of drug use. This was done by repeating the regression techniques described above using P scores from which the effects of items 12, 14, 15, 16 and 18 had been omitted.

Table 6. *Comparison of the P scores of the four drug-using groups*

(Fisher's exact probability test applied to the P scores of the four drug-using groups. Probabilities are as for a one-tailed test.)

	Experimental	Provincial heroin	In-patient injectors	In-patient oral
Experimental	—	$P=0\cdot127$ (split 5/6)	$P=0\cdot304$ (split 5/6)	$P=0\cdot0085$ (split 7/8)
Provincial heroin	—	—	$P=0\cdot571$ (split 4/5)	$P=0\cdot090$ (split 7/8)
In-patient injectors	—	—	—	$P=0\cdot0245$ (split 7/8)
In-patient oral	—	—	—	—

These data were the result of considerable manipulation of the raw data, and it was necessary to combine the four drug-using groups for the difference between the P scores predicted in this way and the corrected P scores actually obtained to reach significance ($t=2\cdot21$; d.f. $=63$; $P<0\cdot05$). Nonetheless, this did show that the elevated P scores among drug-users could not be accounted for simply by the combined effects of the two possible sources of artifact discussed.

Differences between groups of drug-users

Tables 6, 7 and 8 show the differences between the four drug-using groups on P, E and N scores.

Table 7. *Comparison of the E scores of the four drug-using groups*

(t tests applied to the E scores of the four drug-using groups.)

	Experimental	Provincial heroin	In-patient injectors	In-patient oral
Experimental	—	$t=3\cdot24$ ($P<0\cdot01$)	$t=1\cdot51$ (n.s.)	$t=2\cdot86$ ($P<0\cdot01$)
Provincial heroin	—	—	$t=2\cdot00$ ($P=0\cdot06$)	$t=<1$ (n.s.)
In-patient injectors	—	—	—	$t=1\cdot71$ ($P<0\cdot1$)
In-patient oral	—	—	—	—

Table 8. *Comparison of the N scores of the four drug-using groups*

(t tests applied to the N scores of the four drug-using groups.)

	Experimental	Provincial heroin	In-patient injectors	In-patient oral
Experimental	—	$t=1\cdot93$ ($P<0\cdot1$)	$t<1$ (n.s.)	$t=2\cdot44$ ($P<0\cdot05$)
Provincial heroin	—	—	$t=2\cdot66$ ($P<0\cdot02$)	$t<1$ (n.s.)
In-patient injectors	—	—	—	$t=3\cdot29$ ($P<0\cdot01$)
In-patient oral	—	—	—	—

DISCUSSION

It would appear that, even after the data have been manipulated in a number of ways in an attempt to demonstrate that the elevated P scores of the drug-using groups was a spurious finding, there is evidence to suggest that the drug-using groups do in fact have higher P scores than the control groups. How is this finding to be interpreted?

Little is known of the external correlates of scores on the P scale. Eysenck (1964) reports that a group of male prisoners score as high on P (5·1) as a group of male psychotics (4·8), suggesting that the scores do not necessarily reflect the degree to which 'classical' symptoms of psychosis are present. One obvious source of information which might cast light on the meaning of the high P scores in drug-users is the P items themselves.

Before discussing these, however, one point should be emphasized. This is that the use of a number of drug-using groups, as has been done in this study, should not be taken as suggesting the assumption of a unique 'addictive personality', common to all drug-users. There is ample evidence that drug-users, even on the same drug, are a heterogeneous population, in which a number of 'types' can be recognized (e.g. Hill *et al.*, 1962; Meyer *et al.*, 1965). The differences observed between the groups in this study (Tables 6–8) are further evidence of the heterogeneity of the drug-using population. Despite the existence of this heterogeneity, there may be some aspects of personality, common to a number of types of drug-users, which differentiate them from the normal population, e.g. although Hill *et al.* (1962) demonstrated the existence of several different types of MMPI profiles in a population of drug addicts they studied, all but one were characterized by higher Pd scores than the normal population.

This is obviously not to say that every drug addict will necessarily be high on even these aspects of personality; it is the group mean which is higher than that of the general population and there will generally be some overlap between the two populations. In this respect, it should be noted that, for most of the P items which significantly differentiated drug-users from non-users, only a minority of the drug-users actually answered the items in the P direction.

Turning now to the items themselves (Table 3), it can be seen that there is greater homogeneity in the items which do *not* differentiate the two groups from those which do; items 1–7 could all be described as 'paranoid', while there appear to be no items of this type among the items that do differentiate the groups. The latter seem to fall into a group of three items relating to parents (items 13, 19 and 20), two to poor health (items 12, 15) two to having had a hard time out of life (items 14, 16), leaving four items, each of which does not seem to have much obvious connexion with the other items.

Several hypotheses could be advanced to explain the increased frequency of drug-users endorsing these items.

One hypothesis would be that drug-users in reality differ from most people in the directions indicated by the questions, i.e. their fathers were really not good men, etc.

Another hypothesis would be that drug-users do not differ from most people in the directions indicated by the questions, but are simply more 'honest' in their reporting on the questionnaire, i.e. less affected by social desirability.

A third hypothesis would be that drug-users are the type of people who actively choose out 'socially undesirable' responses, both in terms of their drug use and in terms of their endorsement of questionnaire items with low social desirability value.

In relation to the second and third hypotheses, the importance of social desirability in determining scores on P is suggested by the finding, noted earlier, that P scores are higher when the questionnaires are completed anonymously than when the respondent fills in his name. Freedman & Doob (1968) have demonstrated that normal subjects, made to feel deviant by an experimental manipulation, are less susceptible to the social desirability response set than subjects not made to feel deviant. It is probable that groups such as drug-users, criminals, and psychotics do feel different from most people, and this could affect their tendency to respond in a socially acceptable way. It is intended to examine these hypotheses further in later studies.

In terms of extraversion and neuroticism, three of the drug-using groups are significantly high on neuroticism, the fourth approaching significance in the same direction. Two are significantly low on extraversion, the other two not differing significantly from the mean. The only published data on extraversion and neuroticism in drug-users known to the authors are those of Halstead & Neal (1968), who found that their group of British heroin-and-cocaine addicts were significantly high on neuroticism and low on extraversion, and Rosenberg (1969), who found that the mixed group of Australian drug-users he studied were significantly high on neuroticism but average on extraversion. From this small number of studies, it seems that high neuroticism scores are a more consistent feature of drug-users than deviation on extraversion.

The authors are indebted to Dr S. B. G. Eysenck for help and encouragement, and to Dr P. H. Connell, Director of the Drug Dependence Clinical Research and Treatment Unit, for encouragement and permission to study his patients.

REFERENCES

EYSENCK, H. J. (1964). *Crime and Personality*. London: Paladin, 1970.

EYSENCK, H. J. & EYSENCK, S. B. G. (1964). *Manual of the Eysenck Personality Inventory*. London: University of London Press.

EYSENCK, H. J. & EYSENCK, S. B. G. (1968*a*). A factorial study of psychoticism as a dimension of personality. *Multivar. behav. Res.* All-Clinical Special Issue, 15–31.

EYSENCK, H. J. & EYSENCK, S. B. G. (1968*b*). The measurement of psychoticism: a study of factor stability and reliability. *Br. J. soc. clin. Psychol.* **7**, 286–294.

EYSENCK, S. B. G. & EYSENCK, H. J. (1969). Scores on three personality variables as a function of age, sex and social class. *Br. J. soc. clin. Psychol.* **8**, 69–76.

FREEDMAN, J. L. & DOOB, A. N. (1968). *Deviancy: The Psychology of Being Different*. New York and London: Academic Press.

HALSTEAD, H. & NEAL, C. D. (1968). Intelligence and personality in drug addicts: a pilot study. *Br. J. Addict.* **63**, 237–240.

HILL, H. E., HAERTZEN, C. A. & DAVIS, H. (1962). An MMPI factor analytic study of alcoholics, narcotic addicts and criminals. *Q.J. Stud. Alcohol* **23**, 411–431.

KOSVINER, A. *et al.* (1968). Heroin use in a provincial town. *Lancet* i, 1189–1192.

MEYER, A. S., BROTMAN, R. & FREEMAN, A. M. (1965). *Continuities and Discontinuities in the Process of Patient Care for Narcotic Addicts.* U.S. Health Research Council.

ROSENBERG, C. M. (1969). Young drug addicts: background and personality. *J. nerv. ment. Dis.* **148**, 65–73.

SEGRAVES, R. T. (1969). Adrenocortical activity, personality, and body-build. (Unpublished Ph.D. thesis, University of London.)

Manuscript received 16 *February* 1970

Revised manuscript received 29 *May* 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 172–180
With 1 text-figure
Printed in Great Britain

Some Comments on the Nature of Delinquents' Identification with Television Heroes, Fathers and Best Friends

By GRANT NOBLE

Centre for Mass Communication Research, University of Leicester

The minimum context form of repertory grid test was used to measure perceived similarities of both self and ideal-self with father, best friends and male television heroes. The same procedures were adopted for a sample of delinquent boys as for a sample of matched non-delinquent boys. Delinquent boys did not perceive themselves as more like, nor wished to be more like, television heroes than control boys. Delinquents perceived themselves as less like their fathers and best friends and wanted to be less like their fathers and best friends than control boys. Since delinquents identified less with their fathers and best friends than controls and yet identified as much as controls with television heroes, it is suggested that television heroes are thus more salient identification figures for delinquents than for controls. Delinquents, more than controls, thought that television heroes did not act a part, but behaved in the same way both on and off the screen. Finally, it was found that the perceived correlation between ideal-self and real self was lower for delinquents than for controls.

In this study identification is defined as the perceived similarity between images of self and others, experienced in terms of personal constructs. This definition emphasizes identification as a relationship rather than a process, and as a perceived rather than as an actual similarity. Perceptual similarities between self and 'significant others' were assessed by Jones (1961), who used Kelly's repertory grid test. The repertory grid test elicits verbal descriptions of the ways, called constructs, in which the perceiver sees similarities between two of three 'significant other people' and the way in which the third is perceived to differ from the other two. A form of the repertory grid test may then be employed to measure identification of self and ideal-self with 'significant others' and with television heroes.

Glaser (1956) writes that criminal identification 'may occur through positive reference to criminal roles portrayed in mass media' and that 'the family probably is the principal non-criminal reference group'. Glueck & Glueck (1943) found that warm and consistent relationships with parents, but particularly with the father, were the basis for predicting non-delinquency. Masters & Tong (1968) similarly found that recidivists evaluated their fathers less favourably than either non-recidivists or non-offenders. Glaser adds that warm relations inside the family may strengthen the identifications with it. Empirical work by Halloran *et al.* (1969) has suggested that male delinquents over-identify with television heroes, perhaps because there is a lack of adequate male identification models in their environment. This suggestion was put to the test in the present study. If delinquent boys over-identified with television heroes it was thought that they would be more likely than boys in a control sample to think that television heroes were 'real'. Thus it would appear to delinquent boys that heroes were not just playing a part in television portrayals but also behaved in the same way when off the screen.

Both Butler (1960) and Dymond (1954) have found that patients undergoing psychotherapy experienced greater conflict between self and ideal-self than did controls. At a more societal level Merton (1957) has suggested that deviant behaviour may occur 'when there is an acute disjunction between cultural norms and goals and the socially structured capacities of members of the group to act in accord with them'. It would seem reasonable to have expected that delinquent boys would experience greater conflict between self and ideal-self than would boys in a control sample.

Sample. A control sample of non-delinquents was matched in terms of age, sex and intelligence to a sample of delinquent boys. Thirty-nine delinquent boys, aged 12 to 15 years inclusive, in a classifying school* were interviewed, but only after a week's residence, by which time they had 'settled down'. Boys in the control sample would have viewed television in family situations, hence delinquent boys were selected from a classifying school since they would also have recently viewed television in family situations. Boys in the control sample were selected from the bottom streams of a large male secondary modern school, so that both age and intelligence matched that of delinquent boys. Table 1 shows that such matching was tolerably effective. No accurate data were available on record cards concerning social class. From available data it appeared to the investigator that only two delinquent boys and two boys in the control sample had fathers with white-collar occupations. Boys in both samples were from medium-sized urban conurbations.

Table 1. *Matching data for the delinquent and control samples*

(Neither *t* test is statistically significant.)

		Age		Intelligence quotient	
Sample	*n*	Mean	S.D.	Mean	S.D.
Control	30	13·897	1·058	88·333	11·758
Delinquent*	39	13·989	1·307	91·846	7·264
Test		0·313		1·191	

* These data were kindly made available by the psychologists at the classifying school.

METHOD

Procedure. Identification was assessed by means of the repertory grid test. Two of the 'figures' (columns) were Me—now (self) and As I would like to be (ideal-self). Both self and ideal-self ratings could therefore be compared by repertory grid matching procedures with ratings of other figures which included both fathers and best friends and male television heroes. Each boy was interviewed individually by the investigator and importance was attached to establishing rapport, particularly with the delinquent boys. It was explained to the boys that the interview was not a test, in no way affected their classification, and that they could say what they liked without concern. In the event only one delinquent boy proved uncooperative and was excluded from the sample.

Each boy was presented with photographs of The Saint, Steed (Avengers), Sergeant Stone (Z Cars), Dr Who, The Virginian and Big Breadwinner Hog. The selection thus included heroes who broke the law, heroes who enforced the law, a hero acting in a historical setting and a hero participating in fantastic adventures in space and time. Boys were asked to name each of these heroes. When a boy was unable to name a character correctly, a hero of like type was substituted (e.g. Detective Inspector Barlow for Sergeant Stone). If the substitute hero was not recognized, that hero was excluded from the test. Previous research (Halloran *et al.*, 1969) has indicated that delinquent boys most enjoyed viewing exciting aggressive programmes. Discussion with both delinquent and control subjects revealed that the sampled programmes were amongst their favourites, with the possible exception of Dr Who. Four triads (each containing three photographs) were presented, in the same predetermined order, to each boy, who was asked 'in what way are two of these three alike, yet different from the third?'. The resulting constructs were recorded on the repertory grid by the investigator. Each construct consisted of an emergent pole (the perceived similarity between two heroes) and a contrast pole (the perceived dissimilarity of the third hero). Emergent poles were recorded on the left side of the grid and the corresponding contrast pole was recorded directly opposite on the right-hand side of the grid, as can be seen in Fig. 1.

---

* These boys were *en route* to an approved school. The average duration of stay in the classifying school was three weeks.

Constructs: Emergent pole	Figures: Television Heroes, Father, Self and Best Friends										Constructs: Contrast pole
I. *Constructs elicited during sorting procedures*	Dr. Who	Father David	Steed (Avengers)	Best Friend Georgi	Saint	As I would Like to be	Barlow Softly, Softly	Next Best Friend Malcolm	Virginian	Me–now	
Have guns			✓	✓	✓	✓	✓	✓	✓		Never use guns
Plain clothes		✓		✓	✓	✓	✓	✓		✓	Stage clothes
Good fighters			✓	✓	✓	✓	✓		✓		Not such good fighters
Likes best			✓	✓	✓	✓	✓	✓	✓ A		Not like so much
Always fighting			✓	✓	✓	✓	✓		✓	✓	Does not fight so much
Waits for criminals and shoots			✓	✓ B		✓	✓		✓		Does not shoot criminals
II. *Constructs supplied by the investigator*											
Acted		☒	✓	☒	✓	☒	C	☒	✓	☒	For real
Like			✓	✓	✓	✓	✓	✓	✓	✓	Dislike
Tough		✓	✓	✓	✓	✓	✓	✓	✓	✓	Weak
Lifelike		☒	✓	☒	✓	☒	✓	☒	✓	☒	Not lifelike
Starts fights	✓	✓ D	✓	✓	✓	✓	✓				Does not start fights
Does good things			✓	✓	✓	✓	✓		✓		Does bad things
Fast		✓	✓		✓	✓	✓		✓	✓	Slow

A: Likes best of all. B: Georgi waits for me with a shotgun. I wear a tin vest and (he) shoots at me when I come round the corner. C: He wrote 'Softly Softly' himself and he is a policeman when he is not on the telly. D: Well he hits me (confirmed from record sheet).

Fig. 1. The repertory grid test completed by a 12-year-old delinquent boy

Five cards, on which were respectively written, Father, Best Friend, Next Best Friend, Me—Now and As I would like to be, were presented to each boy. The Christian names of the first three people were elicited and written on the cards. Where the boy had no father, either step-father or uncle was substituted. In case the boy could not read, the four triads (each of the three cards representing significant male others) were also presented verbally to each boy. Constructs were elicited and recorded as described above. Six triads (cards and photographs in sets of three) containing both television heroes and people known to the child including self, were then presented and the resultant constructs recorded. The construct 'on television' – 'not on television' was not recorded.

Boys frequently repeated constructs or could not answer, even after prompting; thus the average number of constructs elicited was 5·37 for the control sample and 5·92 for the delinquent sample ($t = 1{\cdot}16$, insignificant). Seven additional constructs were provided if they had not occurred spontaneously, namely like – dislike, tough – weak, acted – for real, starts fights – doesn't start fights, lifelike – not lifelike, does good things – does bad things and fast – slow. Boys were asked what these words meant and if they did not know equivalent sets of words were provided. The choice of additional constructs was determined from a free discussion with delinquent boys about television heroes, the dimensions of meaning (Osgood *et al.*, 1957) and were specifically selected to test the 'reality' hypothesis outlined in the introduction. The first figure was then verbally assessed on every construct in the grid. Each boy had to choose either the emergent pole, the contrast pole or neither, and his answers were recorded by the author as a tick, blank and question mark respectively. The photographs or cards representing the figures in the repertory grid test were assessed one at a time and in the same predetermined order. Television heroes and people known to the child, including self, were alternately presented. Boys were not asked whether he, his father or best friends acted or were for real, nor whether lifelike or not. An example of a completed repertory grid is shown in Fig. 1, and it can be seen that much qualitative data was obtained by use of this technique. This 12-year-old delinquent boy started five fights before being interviewed, and he classified both television heroes and his friends in terms of whether they had a gun or not (construct 1 being very closely associated with constructs 3, 4, 5 and 6: 35 matching scores out of a maximum of 40). One of this boy's pastimes was to dress in a tin vest at which his best friend took pot-shots with a shotgun. This boy above all wanted a gun; a desire which does not bode well for the future.

Scoring procedures. As one of the hypotheses suggested that delinquents would identify more often than controls with television heroes, two scores were calculated for each boy. Each subject's assessment of self was compared with each subject's assessment of the television heroes by repertory-grid matching techniques. The ticks, blanks and queries which appeared in the figure column 'me' were compared with every one of the television hero-figure columns in turn—a maximum total of six comparisons. The number of horizontal correspondences between cells was noted. Two ticks, two blanks or two question marks represented a match. One mark was given for corresponding ticks, blanks or question marks and no marks for any combination of ticks with blanks or question marks. For Fig. 1 the scores were respectively 'me' with Dr Who 5, Steed 4, Saint 6, Barlow 5 and Virginian 5. As the boys had produced varying numbers of constructs, the total number of matching scores between 'me' and all television heroes was divided by the maximum possible (for Fig. 1 the number of constructs available for comparison with 'me' was 11, the number of television heroes was 5, so the maximum score was 55) so that comparison could be made between samples. Matching scores therefore ranged between 0·00 (completely unlike the character concerned) and 1·00 (completely like the character concerned). A score of 0·50 thus represented no match, as did the boy's total matching score in Fig. 1 which was 25/55, namely 0·45. This procedure was repeated for ideal-self assessments by comparing the figure column 'As I would like to be' with every one of the television hero-figure columns in turn. For Fig. 1 the matching scores between 'As I would like to be' and television heroes were respectively Dr Who 1, Steed 10, Saint 10, Barlow 11 and Virginian 9 and the total matching score was 41/55, namely 0·75, which indicated that the boy wished to be like television heroes in many respects.

Two additional scores were calculated to test the hypothesis that delinquent boys were less able than controls to identify with their fathers and best friends. A matching score between self-assessment (me) and father, best friend and next best friend was calculated for each boy using the matching procedures described above. For Fig. 1 the matching scores were respectively 8/11, 4/11 and 7/11, a total of 19/33 (0·58). A matching score between ideal-self assessment (As I would like to be) and father, best friend and next best friend was similarly calculated for each boy, and were respectively for Fig. 1, 4/11, 10/11 and 5/11, a total of 19/33 (0·58). It was expected that delinquent boys, more often than control boys, would over-identify with television heroes and under-identify with fathers and best friends. Scores representing perceived identification of both self and ideal-self with father and best friends and with television heroes were compared between delinquents and controls by *t* tests.*

In order to discover whether television heroes or father and best friends were perceived as more salient identification figures, each boy's matching score for self with television heroes was subtracted from the matching score for self with father and best friends. The resultant difference scores (for Fig. 1 the difference score was 0·58 − 0·45 = 0·13) were then compared between delinquents and controls by the *t* test. The same procedure (ideal-self with father and best friends minus ideal-self with television heroes: Fig. 1, 0·58 − 0·75 = − 0·17) was followed for ideal-self matching scores.

A hypothesis that delinquents, more often than controls, would think television heroes 'real' was also tested. Assessments of the perceived reality of television heroes were calculated by counting the number of times television heroes were said to be for real rather than acted (like in a school play) and lifelike (could meet someone like that) rather than not lifelike (too farfetched). For Fig. 1 the perceived reality score was 6/10. Delinquent scores were then compared with control group scores. It was further expected that delinquents, more often than controls, would perceive differences between self and ideal-self images. The matching scores, between real (me) and ideal-self (As I would like to be), for delinquents (in Fig. 1: 5/11 = 0·45) were compared with the similar matching scores for controls.

Jones (1961) reports that the procedures so far described are reliable. Reliability coefficients were thus not calculated. The validity of the instrument remains difficult to establish. It was thought that boys in the control sample would perceive themselves as more similar to fathers and best friends than to television heroes. This proved to be the case. The matching scores with best friend (0·68), next best friend (0·67), and father (0·66) were considerably higher than the matching scores with television heroes which were at approximately chance or no match levels—Steed (0·53), Sergeant Stone (0·52) and Hog (0·44). Kagan (1958) maintains that identification with the father will decrease with age. The data support such a view, as the matching scores in the control sample for ideal-self with father were 0·79 at 12 years, 0·67 at 13 years, 0·62 at 14 years and 0·59 at 15 years.

RESULTS

The central suggestion that delinquents more often than controls would over-identify with television heroes was unsubstantiated. Table 2 shows that delinquent boys neither perceived themselves more like television heroes, nor wished to be more like television heroes, than did boys in the control sample. The hypothesis that there is a lack of suitable male identification figures for delinquent boys received support. Delinquent boys more often than controls both perceived themselves less like fathers and best friends and wished to be even less like such significant males. Whereas boys in the control sample wanted to be more like fathers and best friends than television heroes, the converse was true of delinquent boys. When a comparison was made between samples of the ideal-self matching scores with fathers and best friends, minus ideal-self matching scores with television heroes, delinquent scores were found significantly different from control scores.

The hypothesis that the perceived correlation between ideal and real self would

---

* Lindquist (1953) suggests that analysis of variance and *t* tests can be applied even when non-parametric measurement techniques have been employed.

Table 2. *Summary table of results comparing identification matching scores, self–ideal-self matching scores and reality scores between controls and delinquents*

(Matching scores have a range of 0·00 (completely unlike characters concerned) to 1·00 (completely like characters concerned). A matching score of 0·50 is indicative of no match.)

	Self-identification matching scores			Ideal-self identification matching scores				
Sample	(*a*) like people	(*b*) like TV heroes	$a-b$	(*c*) like people	(*d*) like TV heroes	$c-d$	Matching scores between ideal self and real self	Proportions of times television characters perceived as real rather than acted and lifelike rather than not lifelike
Mean								
Control	0·675	0·487	0·188	0·663	0·629	0·034	0·630	0·310
Delinquent	0·606	0·469	0·137	0·564	0·630	−0·066	0·561	0·470
t	2·249*	0·487	1·401	2·811**	0·000	2·380*	2·029*	3·780**

* Significant at the 5 per cent level. ** Significant at the 1 per cent level.

Lindquist (1953) recommends that even when marked heterogeneity does exist, it would be better to set a stricter level of significance than abandon the use of the F and t tests. Although directions of difference were postulated in the present study, only values greater than two-tailed levels of significance were accepted.

be less for delinquents than for controls was substantiated. Table 2 shows that delinquents perceived a greater discrepancy between real and ideal-self than did boys in the control sample.

With regard to the perceived reality of television heroes, delinquents significantly more often than controls perceived television heroes as 'real' and lifelike rather than as acted and not lifelike.

DISCUSSION

Cronbach (1956) has pointed out that the repertory grid test 'is indirect and exceptionally flexible, but has serious technical difficulties. The complexity of the data has lured Kelly's students into analyses so involved as to obscure errors in reasoning'. An attempt was made in the present study to standardize the presented figures. Inevitably the whole range of television heroes known to the child was not sampled. The television heroes which became figures for the repertory grid test could have been chosen by the subjects. The author felt that such a procedure might indeed have obscured errors in his reasoning. Results must therefore be treated with caution. The adopted procedure attempted to put boys at their ease but the low matching score calculated between self and ideal-self for the delinquent sample may be because delinquent boys wanted to give the impression that they would in future mend their ways. Several of the boys, however, stated that they did not wish to mend their ways and others stated with some conviction that they did not want to get into trouble in the future. Similar criticisms can be made of Butler's findings with patients undergoing psychotherapy, even though there are adequate theoretical reasons why a discrepancy between self and ideal-self should be apparent amongst 'deviant' individuals. It must also be remembered that a legal definition of delinquency was accepted for the purposes of the present study. If Goffman (1963) is correct this fact alone could have adversely affected these boys' images of self and ideal-self. The legal definition of delinquency covers a heterogeneous range of deviant behaviour; for example, two truancy cases were included in the sampled delinquents. In spite of the heterogeneous nature of the sampled delinquents, the results suggest that delinquent boys perceive themselves less like available male identification figures than do controls. It was further noted that there were negligible differences between delinquents and controls with regard to the way television heroes were perceived as identification figures. However, boys in the control sample wanted to be more like significant male others, while delinquent boys wanted to be more like television heroes. These findings lend support to Glaser's hypothesis that criminal identification may occur through positive reference to criminal roles portrayed on the media when warm relationships inside the family do not strengthen the identification with it.

It is interesting to note in this context that delinquents more often than controls did not perceive themselves like their best friends nor wish to be like their best friends. The best friends of delinquent boys were not always those with whom delinquent acts had been committed. Had this been the case delinquents might have wished to give the impression that they would no longer behave like their best friends. It may be that boys lose, or alternatively never acquire, the ability to identify with people known to them, even with best friends, if they have experienced family situations which do not allow identification within the family. Control boys were more able than delinquents to identify with both fathers and best friends. It seems reasonable to suggest that the greater liking of probationers for television heroes found by Halloran *et al.* occurred because they could not, nor wished to, identify with either their fathers or best friends. Although this suggestion cannot be proved by these data, this author is convinced after reading case histories recording up to twelve years of life in inadequate families that the 'chicken and egg' relationship stated above is correct.

Yet in spite of their background, delinquent boys identified both self and ideal-self more with law-enforcing heroes (Sergeant Stone 0·62: ideal-self) than with law-breaking heroes (Hog 0·50). The corresponding figures for the control sample were 0·62 and 0·40. One cannot therefore simply ascribe deviancy to, or even correlate it with, the influence of criminal television models. There are clearly a large number of other more important factors associated with delinquency. It was noted that delinquents no more than controls perceived considerable differences between self (me) and television heroes. In both samples the figures did not deviate from levels representing no match; neither did delinquents wish to be more like television heroes than controls. The most marked difference between delinquents and controls was that delinquents did not wish to be like their fathers or best friends. Differences were most apparent for ideal-self concepts. This finding could be perhaps interpreted as a difference in motivation between the two samples. The findings suggest that delinquents less often than controls have acquired an integrated conception of self. If delinquents verbally categorize self and ideal-self as different, then, as Foote (1951) argues, 'these categorizations of experience motivate behaviour through the necessary commitment of individuals to particular concatenations of identity in all situations'.

This author has previously contended (1969) that a confusion of the reality-fantasy of television presentations indicates involvement in such presentations. If delinquent boys more often than controls are deprived of suitable male identification figures they may be more involved than controls with male television heroes. The findings suggest that delinquents more than controls were involved with male television heroes since delinquents more often than controls denied that television heroes were merely acting their parts. Thus, while delinquent boys may be relatively unable to identify with people in their immediate environment, delinquent boys identified with television heroes as capably as did controls. Such results, coupled with the finding that delinquents identified more with law-enforcing heroes than with law-breaking heroes, can perhaps be interpreted in the light of Katz & Foulkes' (1962) argument. They write that 'the media may be both sought and used not only to compensate for abortive or ineffective social relations but also to maintain extant and presumably effective ones. It may also be . . . that the mass media behaviour of the alienated and deprived also contains an element of striving to re-establish effective interpersonal contacts.' It may be that delinquent boys interact, albeit only one-way interaction, with law-upholding television heroes in the hope that they may one day be able to interact successfully with individuals whom they can respect.

The author wishes to thank Dr M. S. Hoghughi, Mr A. R. Forrest and Mr N. S. Curtis for their assistance in obtaining samples.

REFERENCES

BUTLER, J. M. (1960). Self concept change in psychotherapy. *Counseling Center Discussion Papers*, Chicago, **6**, 13.

CRONBACH, L. J. (1956). Assessment of individual differences. *Ann. Rev. Psychol.* **7**, 173–176.

DYMOND, R. C. (1954). Effects of psychotherapy on self consistency. *J. counsel. Psychol.* **4**, 15–22.

FOOTE, N. N. (1951). Identification as a basis for a theory of motivation. *Am. sociol. Rev.* **16**, 14–21.

GLASER, D. (1956). Criminality theories and behavioural images. *Am. J. Sociol.* **61**, 433–444.

GLUECK, S. & GLUECK, E. (1943). *Criminal Careers in Retrospect.* New York: Commonwealth Fund.

GOFFMAN, E. (1963). *Stigma: Notes on the Management of Spoiled Identity.* Englewood Cliffs, N.J.: Prentice-Hall.

HALLORAN, J. D., BROWN, R. L. & CHANEY, D. (1969). Mass media-delinquency project. In *Television Research Committee, Second Progress Report and Recommendations.* Leicester: Leicester University Press.

JONES, R. E. (1961). Identification in terms of personal constructs: reconciling a paradox in theory. *J. consult. Psychol.* **25**, 276.

KAGAN, J. (1958). The concept of identification. *Psychol. Rev.* **65**, 296–305.

KATZ, E. & FOULKES, D. (1962). On the uses of mass media as escape: clarification of concept. *Publ. Opin. Q.* **26**, 377–388.

LINDQUIST, E. F. (1953). *Design and Analysis of Experiments in Psychology and Education.* Boston: Houghton Mifflin.

MASTERS, F. G. & TONG, J. E. (1968). The semantic differential test with Borstal subjects. *Br. J. Criminol.* **8**, 20–31.

MERTON, R. K. (1957). *Social Theory and Social Structure.* New York: Free Press.

NOBLE, G. (1969). Identification of the roles of television characters: a process leading to involvement in a television programme. *Bull. Br. psychol. Soc.* **22**, 145. (Abstract.)

OSGOOD, C. E., SUCI, G. J. & TANNENBAUM, T. M. (1957). *The Measurement of Meaning.* Urbana, Ill.: University of Illinois Press.

Manuscript received 24 February 1970

Revised manuscript received 17 June 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 181–184
Printed in Great Britain

Social Support and Conformity: The Effect of Response Order and Differentiation from the Group

BY VERNON L. ALLEN AND JOHN M. LEVINE*

University of Wisconsin

Several recent investigations (Allen & Levine, 1968, 1969, 1971) have supported Asch's (1951) conclusion that the presence of one person who answers prior to a naive subject, and correctly dissents from an erroneous group consensus, significantly reduces conformity relative to a unanimous group. The present study investigated two unexplored variables of potential theoretical importance in explaining the effectiveness of such social support from a partner: (1) response position of the partner, and (2) nature of the social relationship between the subject and the partner relative to other group members.

In previous studies the social supporter has typically answered fourth and the subject fifth (last) in a five-person group (Allen & Levine, 1968, 1969, 1971). Whether it is crucial to conformity reduction that the social supporter respond *immediately* prior to the subject—and hence after the other group members have established a norm—is open to question. One line of argument would predict that a partner answering fourth will be more effective in reducing conformity than a partner answering first. A social supporter responding fourth appears to have bravely taken a stand against a unanimously opposing group. Such attribution of courage may create a positive impression of the social supporter, thereby increasing the subject's willingness to agree with the supporter in opposition to the group.

A second line of argument would make the opposite prediction—that a partner answering first will be more effective in reducing conformity than a partner answering fourth. Responses given by the late supporter may be perceived as partially or wholly dependent on prior group responses; but the early supporter's responses could not have been affected by the group and therefore seem to represent truly independent judgements. Since one important factor underlying the efficacy of social support is the partner's independent confirmation of the subject's own appraisal of the physical and social world (Allen & Levine, 1971), we would predict that the highly independent supporter responding first will be more effective than the less independent supporter responding fourth.

The second variable investigated in this study concerns the effect on conformity reduction of the social relationship between the partner and the subject. In the typical social support study the partner is not differentiated from other group members, i.e. the subject neither knows the identity of the partner nor has interacted socially with him. Yet even a brief social interaction between the subject and partner might create a minimal social relationship that could influence the subject's responses in the group pressure situation, e.g. by decreased feelings of anonymity *vis-à-vis* the partner or by increased feelings of obligation and accountability toward the partner.

In summary, the present experiment investigated the influence of two variables on conformity reduction in the social support situation: the partner's response position, and the social relationship between the subject and the partner. It is hypothesized that greater reduction in conformity will occur when: (1) the partner answers in the first position, and (2) a social relationship exists between the subject and the partner.

A repeated measures analysis of variance design was used. The between-subjects factors were: conditions (unanimous, social support–fourth, social support–first), prior social contact between the subject and the partner (Contact, No-Contact), and sex of the subject (male, female). The within-subjects factor was type of item (visual, information, opinion).

In the unanimous condition the four (simulated) subjects, who answered in the first four positions, agreed on each item. In the two social support conditions, one simulated subject, answering either first or fourth in the group of five, dissented from the erroneous responses of the three other simulated subjects on critical trials by giving popular or correct answers.

---

* Now at the University of Pittsburgh.

In the no-contact conditions, five naive subjects were taken together directly from the waiting room to the experimental room; subjects in these conditions did not meet or interact socially with one another. In the contact conditions, each of four naive subjects interacted briefly, prior to the experiment, with a confederate posing as another subject. In the contact conditions each subject spent about 5 min. in a research room alone with the confederate while both filled out personality inventories. In the social support–fourth and unanimous conditions, the subject was told that the confederate would answer fourth in the five-person group; in the social support–first condition, the subject learned that the confederate would respond first. In all conditions the subject was told to answer last (fifth).

Subjects were 255 undergraduates (123 males and 132 females), enrolled in an introductory psychology course at the University of Wisconsin. Subjects were tested in same-sex groups of five. The apparatus was a Crutchfield-type electrical signalling device (Crutchfield, 1955), consisting essentially of five adjacent booths containing signal lights and answer switches, and a master control panel in an adjoining room. The stimulus series was comprised of 30 items (10 each of visual, information and opinion) drawn from a longer series developed by Tuddenham *et al.* (1956). Each item could be answered by choosing one of the nine alternative answers provided. Of the 30 items used, 12 (four of each type) received group pressure; the remaining 18 items were non-pressure trials.

For each subject a mean conformity score was calculated separately for visual, information and opinion items. Mean conformity scores were obtained by summing the algebraic difference between initial responses given in class several weeks earlier and responses given in the group situation, and dividing by the number of items used. An analysis of variance was conducted on mean conformity scores. Results showed that both the conditions and items main effects were significant at less than the 0·01 level ($F = 36{\cdot}33$ and $F = 19{\cdot}18$, respectively). Moreover, the conditions × contact ($F = 3{\cdot}40$) and items × sex of subject ($F = 5{\cdot}16$) interactions were significant at less than the 0·05 and 0·01 levels, respectively. (The items main effect and the items × sex of subject interaction are not of direct interest in this study and will not be discussed.)

Table 1. *Mean conformity as a function of social contact with the partner and partner's order of response*

	Social contact	No contact	Mean
Unanimous group	0·99	0·82	0·90
Social support–fourth	0·36	0·54	0·45
Social support–first	0·25	0·37	0·31

Presented in Table 1 are mean conformity scores, combined over type of item, for the six experimental conditions. Subsequent tests on overall condition means indicated that conformity in both the social support–fourth (0·45) and social support–first (0·31) conditions was significantly lower than in the unanimous (0·90) condition ($t = 6{\cdot}89$, $P < 0{\cdot}01$, and $t = 8{\cdot}97$, $P < 0{\cdot}01$, respectively). In addition, social support in the first position produced a significantly greater decrease in conformity than did social support given later (0·31 *v.* 0·45; $t = 2{\cdot}08$, $P < 0{\cdot}05$).

Regarding the significant conditions × contact interaction, it is clear from Table 1 that the effect of social contact on conformity varied as a function of degree of group unanimity. In the unanimous condition subjects who interacted with the confederate conformed more than did those who had no contact (0·99 *v.* 0·82). However, in both the social support–fourth and the social support–first conditions, conformity was lower in the contact than in the no-contact conditions (0·36 *v.* 0·54, and 0·25 *v.* 0·37).

Ratings of the social supporter on five 12-point evaluative scales were obtained in the post-experimental questionnaire. For each scale a separate analysis of variance was conducted using two factors: conditions (social support–fourth, social support–first) and social contact (contact, no-contact). No significant effects were obtained on ratings of the social supporter's intelligence, accuracy or independence. But on scales assessing liking significant differences were found. Significantly greater liking of the partner was attributed to the group when the partner answered first rather than fourth ($F=3{\cdot}91$; $P<0{\cdot}05$). Surprisingly, subjects expressed greater personal liking for the supporter ($F=3{\cdot}41$; $P<0{\cdot}10$) and estimated that the group liked the supporter more ($F=7{\cdot}21$; $P<0{\cdot}01$), in the no-contact than in the contact condition.

In sum, results of the present study showed that conformity was significantly lower when a social supporter (partner) answered first rather than fourth (just prior to the subject). A significant conditions × social contact interaction indicated that, in the unanimous condition, prior social contact between the subject and confederate increased conformity; however, in the social support conditions, social interaction with the confederate (partner) decreased conformity.

That conformity to the group was lower when a dissenting partner answered first rather than fourth is consistent with our understanding of one of the mechanisms which may underlie the social supporter's effectiveness in reducing conformity. Specifically, it is suggested that the partner's correct or popular answer provides an independent source of information that confirms the subject's own perception of the physical and social world (Allen & Levine, 1971). The order in which the partner responds should be one important determinant of his perceived adequacy as an independent source of information. Although a partner responding after other group members might be perceived as partially influenced by the group's responses, a partner responding first in order would probably be perceived as independently appraising the stimulus. Thus we expected (and obtained) more conformity reduction in the social support–first than in the social support–fourth condition.

Recall that we found opposite results in the social support and unanimous conditions concerning the effect on conformity of prior social contact between the subject and confederate. In the two social support conditions conformity was lower when the subject and confederate (partner) had prior social contact; in the unanimous condition conformity was slightly higher in the contact than in the no-contact condition. In order to explain these results we must recognize that two closely related psychological mechanisms—anonymity and differentiation—probably are operating in the social contact situation.

An explanation based on anonymity assumes that, as a result of prior social contact with the confederate, the subject perceived that one other group member could identify his responses in the group pressure situation. Data on the relationship between publicness of response and conformity (Asch, 1956; Deutsch & Gerard, 1955; Levy, 1960) suggest that if a subject's responses are identifiable by another group member, the subject will respond similarly to the observer, e.g. conform if the observer conforms, dissent if the observer dissents.

The explanation based on differentiation assumes that the subject perceived the confederate and himself as a social unit, in contrast to the undifferentiated group. Although social contact between the subject and confederate was brief and superficial, an elementary social relationship may have been created that caused the subject to feel accountable (*vis-à-vis* the confederate) for his behaviour (Festinger *et al.*, 1952; Singer *et al.*, 1965). This feeling of accountability or responsibility, in turn, may have motivated the subject to respond similarly to the confederate.

The research reported herein was performed persuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare, under the provisions of the Cooperative Research Program (Center no. C-03, Contract OE 5-10-154).

REFERENCES

ALLEN, V. L. & LEVINE, J. M. (1968). Social support, dissent and conformity. *Sociometry* **31**, 138–149.

ALLEN, V. L. & LEVINE, J. M. (1969). Consensus and conformity. *J. exp. soc. Psychol.* **5**, 389–399.

ALLEN, V. L. & LEVINE, J. M. (1971). Social support and conformity: the role of independent assessment of reality. *J. exp. soc. Psychol.* **7**, 48–58.

ASCH, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (ed.), *Groups, Leadership, and Men*. Pittsburgh: Carnegie Press.

ASCH, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychol. Monogr.* **70**, no. 9. (Whole no. 416.)

CRUTCHFIELD, R. S. (1955). Conformity and character. *Am. Psychol.* **10**, 191–198.

DEUTSCH, M. & GERARD, H. B. (1955). A study of normative and informational social influences upon individual judgment. *J. abnorm. soc. Psychol.* **51**, 629–636.

FESTINGER, L., PEPITONE, A. & NEWCOMB, T. (1952). Some consequences of de-individuation in a group. *J. abnorm. soc. Psychol.* **47**, 382–389.

LEVY, L. (1960). Studies in conformity behavior: a methodological note. *J. Psychol.* **50**, 39–41.

SINGER, J. E., BRUSH, C. A. & LUBLIN, S. C. (1965). Some aspects of de-individuation: identification and conformity. *J. exp. soc. Psychol.* **1**, 356–378.

TUDDENHAM, R. D., MACBRIDE, P. & ZAHN, V. (1956). Studies in conformity and yielding: I. Development of standard experimental series. (Tech. Report no. 1, ONR, Contract NR 170–159, University of California.)

Manuscript received 6 June 1969

Revised manuscript received 14 July 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 185–186
Printed in Great Britain

Conservatism and Response to Sexual Humour

BY D. R. THOMAS, J. D. SHEA AND R. G. RIGBY

University of Queensland, Australia

In their presentation of a new measure of conservatism, Wilson & Patterson (1968) listed a number of characteristics as being expected in extreme conservatism. These included: preference for conventional art, clothing and institutions and an anti-hedonistic outlook. The Conservatism Scale also contains items of a sexual nature (e.g. 'chastity', 'striptease shows', 'nudist camps'), indicating that conservatism involves some disapproval of sexual behaviour. Thus it would be expected that conservative persons would react less favourably to humour having a sexual theme. The present study examines the relationship between conservatism and reaction to a cartoon having a sexual theme.

A cartoon sequence containing 12 pictures was selected as the stimulus. The nature of the cartoon was such that it was difficult to see the point unless all the pictures were scanned. Only people who definitely stopped to read the cartoon were taken as subjects. Most subjects took between 5 and 10 sec. to read it. The cartoon itself depicted a nude female 'growing' out of a doorbell and the startled response of a male caller.

The cartoon was placed on the outside of a glass pane, next to a side entrance door into the social sciences building at the University of Queensland. This entrance is predominantly used by education, psychology and sociology students going to and from a lecture theatre. The position of the cartoon was such that people coming in the door were likely to see it, while those going out were not likely to see it. The height of the cartoon was adjusted so that the facial expressions of the subjects were clearly visible through the glass pane.

A TV camera was set up in the entrance foyer and focused on the area just above the cartoon. Other equipment was placed around the camera, so it appeared to be temporarily stored in the foyer. The camera, which could not be seen to be in operation from the foyer, was monitored by the experimenters in a nearby storage room. All persons who stopped to look at the cartoon were recorded on video-tape. After a trial run in which 10 subjects were recorded, two of the experimenters waited in the foyer with the conservatism questionnaire and approached all the people who had looked at the cartoon under the pretext of doing 'a random survey of social attitudes of people passing through the building'. Only one person refused to participate, giving lack of time as a reason. In this way 39 subjects, who had both looked at the cartoon and filled in the questionnaire, were obtained. None of the subjects showed any indication of connecting the camera or questionnaire with the cartoon.

A group of five psychology graduate students rated the favourability of the recorded reactions of the subjects to the cartoon on a seven-point scale. The average inter-rater reliability was 0·85 with a range from 0·75 to 0·92. This result is consistent with the findings of Ekman *et al.* (1969) that happiness is the most stable dimension in facial displays of emotion in terms of agreement between observers, and agreement across cultures. Accordingly, the ratings of the five judges were combined to obtain a single index of favourability for each subject.

The product-moment correlation between conservatism and favourability of reaction was −0·44 (d.f. = 37; $P < 0·005$), indicating a significant tendency for conservative subjects to react less favourably to the cartoon. With a sample of 20 males and 19 females, a point biserial correlation of 0·36 (d.f. = 37; $P = 0·01$) between sex and favourability was obtained. Males showed more favourable reactions. As there was also a 0·39 correlation between sex and conservatism (females being more conservative), the partial correlation between favourability and conservatism was computed, with sex partialled out. The resulting correlation of −0·35 indicates that there is a negative relationship between conservatism and favourability independently of sex differences in conservatism.

The results obtained support the validity of the Conservatism Scale as a general measure of conservatism. The idea that humour with a sexual theme is not usually appreciated by conservative' persons is frequently met in informal observation. Also the results support those of Wilson & Patterson (1969), who found that conservative persons rate sexual humour as significantly less funny than liberals. The correlation obtained in their study between conservatism and 'funniness' (−0·24) is lower than that obtained here. This may reflect differences in the cartoons used or a lower validity for questionnaire measures of 'funniness'.

Finally, the present study indicates one way in which personality questionnaires may be validated by the observation of behaviour in non-laboratory settings (cf. Webb *et al.*, 1966).

REFERENCES

EKMAN, P., SORENSON, E. R. & FRIESEN, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science, N.Y.* **164**, 86–88.

WEBB, E. J., CAMPBELL, D. T., SCHWARTZ, R. D. & SECHREST, L. (1966). *Unobtrusive Measures: Nonreactive Research in the Social Sciences.* Chicago: Rand McNally.

WILSON, G. D. & PATTERSON, J. R. (1968). A new measure of conservatism. *Br. J. soc. clin. Psychol.* **7**, 264–269.

WILSON, G. D. & PATTERSON, J. R. (1969). Conservatism as a predictor of humor preferences. *J. consult. clin. Psychol.* **33**, 271–274.

Manuscript received 26 *May* 1970

Revised manuscript received 26 *July* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 187–188
Printed in Great Britain

Ethnocentrism and the Evaluation of Accented Speech

BY HOWARD GILES

University of Bristol

In a previous study using the 'matched-guise' technique (Giles, 1970*a*) it was found that 17-year-old sixth-formers were able to assign social prestige values to various forms of British regional and foreign accented speech. A follow-up study by the same author (1970*b*), using native accented speech samples with 21-year-old college students as subjects, showed a rho correlation of +0·938 for accent prestige between the two studies. Nevertheless, there was a significant tendency for these subjects to rate accents more favourably than their younger counterparts. This differential evaluation by the two age-groups, although admittedly to different stimuli, could be explained in terms of the more diffuse social environment surrounding the maturer subjects. In other words, it could be argued that the social qualities of college life are often such as to mould less ethnocentric and more liberal attitudes towards the characteristics of outgroups.

While Lambert (1967) has reviewed studies showing age, sex, social class, regional and linguistic competence differences in the evaluation of personality content from various dialects and languages, no work has been conducted to ascertain the effect of personality differences between *listeners* in evaluation. On the basis of the above assumption, it was hypothesized that the more ethnocentric an individual's orientation, the less favourable his evaluation of regional accented speech would be. Secondly, it would seem to follow that the highly ethnocentric would react relatively more favourably towards the Southern Standard accent or Received Pronunciation (RP) than the less ethnocentric, simply because of its superior social prestige value in relation to all other regional varieties.

Sixty-three sixth-formers from South Wales and Somerset were administered the British E scale constructed by Warr *et al.* (1967) and an accent evaluation task. From this subject pool, five male and five female subjects from each region were ultimately selected with the highest and the lowest E scores. Thus 20 E+ and 20 E− subjects were derived, matched for age, sex and region. All E+ scores obtained were higher than 83, while all E− scores were below 73, affording group means of 99·0 and 60·3 respectively, with a mean age of 17 years 6 months. After completion of the 24-item ethnocentrism questionnaire, all subjects were required to evaluate the aesthetic, status and communicative contents of six speakers reading the same 35-sec. passage of neutral prose on seven-point rating scales. Thus subjects were requested to rate (*a*) how pleasant–unpleasant they found each voice, (*b*) how much social prestige the voices possessed, and (*c*) how comfortable–uncomfortable they would feel if they were themselves to interact with each of the speakers. The voices were actually produced by the same male speaker assuming the same speech rate, vocal intensity, timbre and personality throughout the six guises, which were considered quite realistic in a previous study (1970*a*). The subjects were familiarized with the stimulus passage before the commencement of the rating procedure. The specific accents used in the study are given in Table 1, together with the *t* values between the two groups of subjects.

From Table 1 it can be seen that the results tend to support both hypotheses. There is a general tendency for the E+ subjects to rate regional accents less favourably over all three evaluative dimensions in relation to E− individuals, thereby displaying less tolerance towards regional variations in accent usage. Moreover, this lack of tolerance towards non-standard speech is accentuated by means of a more favourable reaction to RP than the E− group, particularly with regard to its aesthetic content. However, it appears both groups report to be equally at ease with RP speech. It will also be noted that the extent of evaluative divergence between the two groups is a function of the specific regional accents used. Whether a similar pattern of evaluation would emerge for the personality content of speech, as seems likely, is a matter for further empirical inquiry. Nevertheless, as Ramsay (1968) has shown personality differences to be an important factor with regard to individual variations in the encoding mechanism of speech, the present findings suggest no less a status for personality differences in the decoding process.

Table 1. *t values between E+ and E− subjects*

(A positive value indicates a more favourable mean rating score for E− subjects; d.f. = 38.)

	Evaluative dimensions		
	Aesthetic	Status	Communicative
Irish	+1·930	+1·763	+0·792
Southern Welsh	+1·070	−0·704	+3·996****
Received pronunciation	−2·219*	−1·363	0
Birmingham	+1·674	+1·318	+1·848
Somerset	+2·136*	+2·921***	+2·873***
Northern	+2·579**	+1·665	+2·514**

**** $P<0·001$. *** $P<0·01$. ** $P<0·02$. * $P<0·05$.

The author thanks Mr P. F. Powesland, Bristol University, for his supervision of this project.

REFERENCES

GILES, H. (1970*a*). Evaluative reactions to accents. *Educ. Rev.* **22**, 211–227.

GILES, H. (1970*b*). A repetition study on evaluative reactions to accents. (Unpublished paper, University of Bristol.)

LAMBERT, W. E. (1967). A social psychology of bilingualism. *J. soc. Issues* **23**, 91–109.

RAMSAY, R. W. (1968). Speech patterns and personality. *Lang. & Speech* **11**, 54–63.

WARR, P. B., FAUST, J. & HARRISON, G. J. (1967). A British ethnocentrism scale. *Br. J. soc. clin. Psychol.* **6**, 267–277.

Manuscript received 20 July 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 189–190
Printed in Great Britain

Adaptation of a Hindi Version of the Junior Eysenck Personality Inventory

By B. S. GUPTA

Department of Psychology, Government College, Kurukshetra, India

A Hindi version of the Junior Eysenck Personality Inventory was prepared and administered to 386 boys and 386 girls of eight schools (four boys' and four girls' schools), selected randomly from all the 67 high and higher secondary schools of Ambala and Thanesar tehsils of Ambala and Karnal districts respectively, in the State of Haryana, India. All the children present on the day of testing, in various age-groups, viz. 13+, 14+, 15+ and 16+, in the 8th, 9th, 10th and 11th classes, were studied. Children who left any item of the inventory unanswered were eliminated from the data. The number of such children was 71.

Table 1. *Means and S.D.s for the Extraversion, Neuroticism and Lie scores*

Age	*n*	E		N		L	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
13+							
Boys	120	18·66	3·95	12·41	5·13	3·17	2·07
Girls	120	17·96	3·96	14·80	5·20	3·37	2·36
14+							
Boys	90	18·43	3·88	12·49	5·65	3·21	2·35
Girls	90	18·21	4·85	15·79	5·69	3·22	2·38
15+							
Boys	96	18·24	4·43	12·86	5·99	2·26	2·41
Girls	96	18·03	3·92	15·24	5·68	2·80	2·57
16+							
Boys	80	18·30	4·22	11·80	5·79	2·68	2·21
Girls	80	17·66	4·15	14·92	4·47	2·25	2·01

The results, given in Table 1, revealed no significant differences between boys and girls in respect of extraversion (E) and lie (L) scales, whereas such differences were found to exist in respect of the neuroticism (N) scale. No significant increase or decrease in E, N and L scores was found with increase in age. A further investigation with subjects of 9, 10, 11 and 12 years of age might throw some light on this aspect. A comparison with English norms (Eysenck, 1965) revealed that both the Indian groups were slightly more extraverted and neurotic than the English standardization groups. No such differences were detected in respect of the L scale.

Table 2. *Correlations between scores*

Age	E with N		E with L		N with L	
	Boys	Girls	Boys	Girls	Boys	Girls
13+	−0·24	−0·10	−0·22	−0·24	−0·35	−0·33
14+	−0·18	−0·15	−0·10	−0·13	−0·26	−0·30
15+	−0·22	−0·20	−0·18	−0·19	−0·40	−0·18
16+	−0·20	−0·21	−0·23	−0·26	−0·35	−0·37

Table 3. *Split-half reliability*

	E		N		L	
Age	Boys	Girls	Boys	Girls	Boys	Girls
13+	0·78	0·81	0·84	0·84	0·76	0·72
14+	0·79	0·79	0·84	0·85	0·76	0·75
15+	0·85	0·84	0·86	0·86	0·78	0·77
16+	0·86	0·86	0·87	0·87	0·77	0·77

The correlations between scores on E, N and L, given in Table 2, for different age and sex groups, revealed a slight negative correlation of about −0·19 between E and N. A similar trend was discerned between E and L, and N and L scores, though the correlation between the latter was rather large. The split-half reliabilities (Table 3) ranged between 0·72 and 0·87. Correlations and reliabilities were in good agreement with the English norms.

Thus the data suggest that the Hindi version of JEPI gives results similar to those obtained with the original version in England.

REFERENCES

EYSENCK, S. B. G. (1965). *Manual of the Junior Eysenck Personality Inventory*. London: University of London Press.

Manuscript received 15 *April* 1970

Revised manuscript received 29 *June* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 191–192
Printed in Great Britain

Gibson Spiral Maze and MPI Scores: A Comparison using Normal Adolescent Subjects

BY KENNETH G. McDONALD AND ADRIAN D. PARKER

St George's Hospital, Morpeth, Northumberland

Of the many different possible approaches to the assessment of personality, relatively little attention has been given to techniques such as psychomotor tests giving measures of speed and accuracy of functioning which may reflect personality differences. Nevertheless some groundwork exists. Yates (1960) concluded: (1) that neurotics show no significant disorganization on simple measures, but that they do on more complex tests of dexterity; (2) that there is a postulated difference in psychomotor performance between introverted and extraverted neurotics from the Eysenckian theory of excitation–inhibition imbalance; the growth of reactive inhibition annulling drive and reducing efficiency differently for the two groups. Yates also concluded (p. 54) that psychomotor tests are 'eminently suitable for precise predictions in terms of the theory, particularly those tests involving continuous performance'.

Gibson's Spiral Maze Test (GSM) appears a potentially valuable psychomotor test requiring the subject to trace a path from the centre of a spiral maze free of blind alleys to an exit at the circumference (Gibson, 1965). Subjects are instructed to work as quickly as possible but to avoid touching either obstacles in the path or demarcation lines. The subject is also urged every 15 sec. to work quickly. Time for completion is recorded in seconds (t), and a weighted error score (e) is obtained. Gibson reports that performance relates to teacher-rated degrees of 'naughtiness' and later to delinquency, but concludes (p. 7) that 'it remains to be demonstrated in what way the test is a valid test of personality characteristics'.

It was therefore decided to examine GSM performance in relation to scores on the dimensions of Extraversion and Neuroticism in a non-psychiatric sample. Gibson considered (p. 6) that 'an adjusted e score is probably the most useful single measure of psychomotor competency'. Consequently corrected e scores, as described later, were selected for analysis. Gibson's revised scoring procedure (Gibson, 1969) was not adopted in this study. In addition directly to comparing measures of E and N with GSM performance, the prediction that there would be a significant difference in psychomotor competency between those high N scorers who scored high on E, and those who scored low on E, was set up for evaluation.

The GSM and the short form of the MPI were administered to 50 senior grammar-school pupils (26 girls, 24 boys). Ages ranged from 16 to 19 yr., median age 17·2 yr., thus providing a group relatively homogeneous for age and scholastic aptitude. Non-parametric statistics were used, giving median $E = 7{\cdot}8$, median $N = 7{\cdot}5$. The inter-scorer reliability of error scoring was checked by independent scoring of all maze cards by both authors, and was found to be high ($r = 0{\cdot}9$). Nevertheless, all protocols showing any discrepancy were re-examined, and agreed error scores obtained. The reported relationship between these and t scores for normal populations was confirmed ($r = -0{\cdot}5$). In terms of raw scores, $Q_1e = 7{\cdot}8$, $Q_2e = 14{\cdot}5$, $Q_3e = 21{\cdot}8$, and $Q_1t = 29{\cdot}4$, $Q_2t = 34{\cdot}5$, $Q_3t = 41{\cdot}7$. For further examination it was necessary to obtain corrected error scores (e_c; partialling out the effects of time) by the graphical correction of individual raw scores from regression lines of time and error percentiles. Rank order correlations were then calculated between e_c and E ($r = -0{\cdot}03$; n.s.), and e_c and N ($r = +0{\cdot}37$; $P < 0{\cdot}01$). In order to examine whether there was any difference in psychomotor competency in the way predicted, a group ($n = 31$) scoring at or above the median N score was selected and divided into those scoring above the median for E (NE; $n = 12$) and those scoring below the median for E (NI; $n = 19$). Of the NI group, seven fell above the median e_c score for the total group and 12 fell below. The NE group was divided seven below and five above the median. These groups therefore fail to show any significant difference in psychomotor competency ($\chi^2 < 1$). However, the results show that the time involved in completing the maze is somewhat short for the examination of reactive inhibition effects.

Corrected error scores were therefore found to relate to degree of emotional lability as measured by the short form of the MPI in a group of grammar-school pupils, although no relationship was found with scores on the extraversion scale. Thus evidence was found that maze performance does relate in some way to personality. This is encouraging, but serves as a caveat to the use of such tests purely to measure psychomotor functioning without considering personality variables and emotionality in particular. No significant difference was found in psychomotor competency within the group of those scoring high on N (emotionally labile) according to whether they were classified as introverted or extraverted, probably because the GSM is too short a test for evaluating predictions about reactive inhibition. The statistical procedures necessary to evaluate GSM performance and reach these conclusions confirm the need for a revised scoring procedure.

REFERENCES

GIBSON, H. B. (1965). *Manual of the Gibson Spiral Maze*. London: University of London Press.

GIBSON, H. B. (1969). The Gibson Spiral Maze Test: retest data in relation to behavioural disturbance, personality and physical measures. *Br. J. Psychol.* **60**, 523–528.

YATES, A. (1960). Abnormalities of psychomotor function. In H. J. Eysenck (ed.), *Handbook of Abnormal Psychology*. London: Pitman.

Manuscript received 12 *May* 1970

Br. J. soc. clin. Psychol. (1971), **10**, p. 193
Printed in Great Britain

Extreme Responses of Psychiatric Patients on the Semantic Differential: An Explanation for Discrepant Findings

By H. G. I. LOVE

Porirua Hospital, Porirua, New Zealand

Parsonson (1969) reports a study which apparently fails to support earlier findings by Arthur (1966) that hospitalized psychotic patients give extreme responses significantly more often than neurotics on the same form of the Semantic Differential. Parsonson's (1969) findings are similar to those earlier reported by Love & Priest (1967). The figures from the three studies appear as Table 1, the mean number of extreme responses appearing in the body of the table.

Table 1. *Mean number of extreme responses for neurotic, psychotic and normal subjects in three studies*

	Neurotic	Psychotic	Normal
Arthur (1966)	27·9	41·3	17·1
Love & Priest (1967)	28·01	31·2	23·3
Parsonson (1969)	26·60	33·35	19·40

Venables (1964) has presented considerable data to support the hypothesis that acute patients are suffering a breakdown in sensory input filtering mechanisms, and as a result are bombarded with persistent, intense stimuli. To reduce this painful condition responses become stereotyped as the patient moves from the acute phase to the chronic phase of illness, so that acute patients would be expected to give fewer extreme responses than chronic ones. The data reported in Table 1 are consistent with this hypothesis, both Parsonson (1969) and Love & Priest (1967) using an acute patient population (less than 3 months in hospital), whilst Arthur (1966) used a chronic population (mean time in hospital 54 months). Although the figures in the table clearly give weight to Venables' (1964) position, a more convincing demonstration would come from a longitudinal study. This is currently in progress.

An important concomitant of this is that, as differential diagnosis is of most importance at the acute stage of an illness, the reported findings suggest that the Semantic Differential will not be useful as a diagnostic tool, but, as Parsonson (1969) suggests, may serve as a 'prognostic indicator'.

REFERENCES

Arthur, A. Z. (1966). Response bias in the semantic differential. *Br. J. soc. clin. Psychol.* **5**, 103–107.
Love, H. G. I. & Priest, H. F. (1967). Extreme responding in psychiatric patients. *N.Z. med. J.* **66**, 882–883.
Parsonson, B. S. (1969). Extreme response tendencies on a semantic differential. *Psychol. Rep.* **24**, 571–574.
Venables, P. H. (1964). Input dysfunction in schizophrenia. In B. A. Maher (ed.), *Progress in Experimental Personality Research*, vol. 1. New York: Academic Press.

Manuscript received 28 May 1970

Book Reviews

Put Away: A Sociological Study of Institutions for the Mentally Retarded. By Pauline Morris. (Foreword by Professor Peter Townsend.) London: Routledge & Kegan Paul, 1969. Pp. xxxiii + 355. £3.00.

This, the first detailed study of mental subnormality hospitals in England and Wales, was carried out at the invitation of the National Society for Mentally Handicapped Children. The Society had expressed its concern with the problem by making a research grant to Professor Peter Townsend of the Department of Sociology in the University of Essex. The Ministry of Health co-operated at all stages of the investigation, as did the staff of the 34 hospitals included in the survey, these being a large representative sample of all such institutions.

The book was published within a few months of the appearance of an official report on allegations of ill-treatment of patients in a Welsh subnormality hospital, and the committal to prison of a male nurse for assaulting a patient in an English subnormality institution. In the circumstances it is as well to bear in mind Dr Morris's final sentence: 'The staff in the subnormality hospitals badly need better instruction and guidance, as well as more time to devote to fewer patients, but it should not be forgotten that in one important sense many of them accept those whom society chooses to reject.'

Nevertheless, the picture, soberly drawn and tactfully presented, is not a pretty one. Buildings are Dickensian; food, while ample, is dull and the budget stringent; lack of personal possessions extends to clothes and even toothbrushes; overcrowding in dormitories is such that there is not room for furniture other than beds. Although 83 per cent of all patients are fully ambulant, and only 9 per cent bed-fast, a majority have little to occupy them during the day, and few social relationships. A majority of children have no toys, and only 43 per cent of them attend hospital schools, often on a part-time basis and with inadequately qualified teaching staff.

Many of these facts, important as they are, could be gleaned by a relatively unsophisticated observer and have been the subject of various Ministry of Health circulars; they are probably well appreciated by most staff and some visitors to hospitals. The strength of this book lies, however, in the careful documentation of staff attitudes, and the tensions that exist within these communities which are usually both physically and psychologically isolated from the rest of society. The fragmentation of individual effort, the serious lack of communication between medical and nursing staff, and the suspicion and hostility expressed by them towards specialist services provided by teachers, psychologists, occupational and industrial therapists and physiotherapists make tragic reading.

Despite the explosive increase in research and major advances in the understanding of mental subnormality which occurred during the last two decades, British mental deficiency hospitals clearly remain, with a few outstanding exceptions, wedded to those 19th-century custodial conceptions which were current when the service was initiated. In the reviewers' opinion the reason for this state of affairs is not difficult to determine. The vast majority of the institutionalized mentally subnormal 'patients' are neither physically nor mentally ill. They are, however, cared for within units called hospitals predominantly by members of two professions whose background training is either largely irrelevant or positively inappropriate to the task of shaping their maladaptive behaviour and helping them to make the best use of limited assets. It is small wonder that doctors and nurses feel frustrated and insecure, and suspicious of those who by virtue of their more relevant professional attitudes and skills constitute a threat to their established authority.

Put Away is a book which deserves to be read by all those who care about how an affluent and enlightened society treats its less fortunate members. For professional psychologists it may have the unfortunate effect of diminishing recruitment to a service which badly needs their skills, not only as intelligence testers (although Chapter 4 documents the absence or inadequacy of psychological assessments in hospitals), but also to design and assist in the execution of programmes of education and rehabilitation, as well as in fundamental research.

One further point is perhaps relevant here. Although it was not part of Dr Morris's brief to assess the effects on the patients of the physical and psychological deprivations which she describes, carefully controlled research on this topic has been conducted in the United States with results suggesting the hazards of using institutionalized samples for research purposes as representative of the total subnormal population.

A. M. AND A. D. B. CLARKE

Cross-cultural Studies. Edited by D. R. Price-Williams. (Penguin Modern Psychology Readings.) Harmondsworth: Penguin Books, 1969. Pp. 384. 45p.

This is a collection of 23 papers which have been published elsewhere and brought together as selected readings as part of the series 'Penguin Modern Psychology Readings'. Thirteen of the papers concern work in sub-Saharan Africa, seven are general in nature or compare work from several countries and only three report on work done in a particular area outside Africa. The papers concern ability tests, visual illusions, pictorial depth perception, field dependence, conservation (Piaget), achievement motivation and Freudian hypotheses. Some of the papers could be referred to as psychology from other cultures rather than cross-cultural but most have comparative value. It is not only useful to have the papers brought together in this way but it also provides, in the bibliographies, references to most of the work in this field, depressingly small though the field is. Most of the papers were originally published in the last five years except for those concerning Freudian hypotheses. This presumably reflects the greater interest in those aspects of cross-cultural psychology in the 1950s and the present decline of 'Culture and Personality' studies.

JOHN ORLEY

Smoking, Health and Behavior. Edited by Edgar F. Borgatta and Robert R. Evans. Chicago: Aldine, 1968. Pp. xii+288. No price given.

This book brings together papers and reports presented at the 1967 National Research Conference on Smoking and Health, held at the University of Wisconsin. These papers range over a very wide area; they are grouped in four parts: (1) 'The Health Issue'; (2) 'Social and Psychological Aspects of Smoking'; (3) 'Current Research on Psychological Aspects of Smoking'; (4) 'Implications of the Smoking–Health Problem for Society'.

The 21 papers represented range in quality from moderate to good; this is not unusual in conferences of this kind. It is to be noted that the conclusions reached with respect to the effects of smoking on health are more cautious and less strident than those arrived at by the various Medical Commissions here and in the U.S.A. which have studied the topic. This is in part due to the greater sophistication of modern research methods, which is itself perhaps a function of the criticisms made of the earlier work. Thus attempts are now made to control for the influence of hereditary factors; this is done by comparing identical twins who smoke with twins who do not. This very important work (which seems to be quite unknown in this country) has shown, for instance, that 'extensive comparison of disease and function of the cardiovascular systems of these twins showed no difference between the smoking and non-smoking twins. The incidence of cardiovascular abnormalities was concluded chiefly to be determined by heredity with no discernible effect from smoking'. A greater influence of smoking on health was found in relation to the pathogenesis of chronic broncho-pulmonary disease. It is only by thus ruling out important sources of error and by making the study of the effects of smoking more specific that we can hope to gain proper scientific understanding of the issues involved. This book, by reviewing and emphasizing these points, makes a contribution to the subject.

H. J. EYSENCK

The Nature of Emotion: Selected Readings. By Magda B. Arnold. (Penguin Modern Psychology.) Harmondsworth: Penguin Books, 1968. Pp. 397. 45p.

As the title suggests, the scope of the book is wide, dealing with the subjects from phenomenological, behavioural, biological and physiological points of view. The first three chapters (James, Lehmann and Cannon) deal with the James–Lange theory of emotion. These lead to a section on the instinctive basis of emotional behaviour (Shand, McDougall and Plutchik), followed by a chapter each on the Freudian and Jungian writings on emotion (Rapaport and Hillman, respectively). The next section, 'Emotion as Significant Experience', consists of translations of two turgid and obscure articles by Krueger and Klages respectively and of one by Dumas describing autonomic changes associated with so-called emotional shocks. There follows a section entitled 'Emotion as Hypothetical Construct'. This contains an early, though not the earliest, statement by Duffy of the position of the 'energetics group'. In view of the latter's influence one might have expected a stronger representation here, such as an extract from Lindsley's article in Stevens' *Handbook;* in fact, it leads to an extract from *The Organization of Behaviour:* emotion as incompatibility of phase sequences. The next three chapters (Claparède, Leeper and Pradines) are also concerned with the disruptive aspects of emotion. Next comes 'Emotions as Organized Tendencies'. What possible precise meaning can be given to the term 'tendency'? It is any more meaningful to define emotions as organized tendencies leading to action than to talk of instincts?

The penultimate section is a treatment of the morphology of emotional expression. It includes Spitz's ontogenetic study of expression which Arnold describes (debatably) as complementary to that of Bridge. As the latter is out of print, it would have been helpful to have had extracts from it. Similarly, Michotte's statement on the greater importance of movement over 'frozen' facial expression in the recognition of emotion would have had greater impact had it been preceded by one of the 'frozen' classics such as Piderit or Féléky. How interesting also to have reproduced an extract from Ekman's work on 'micro-expressions'!

The final section on physiological aspects of emotion, as might be expected, is one of the stronger ones, for it is here that solid achievements lie. There are extracts from both Cannon's and Papez's classic studies. A substituted chapter by Delgado might perhaps have been dropped for MacLean's revision of Papez, leading more naturally to Dr Arnold's own chapter on the neural mediation of the emotional components of action.

It is easy to criticize a book such as this for missing out one's own favourites and for putting in what frankly is second-rate work. However, perhaps to omit the second-rate is to give a misleading picture of the field. For, as Dr Arnold herself says, emotion 'is one of the topics that have immediate appeal and human relevance, but yield only reluctantly to scientific research and explanation'.

JOHN HUTT

Motivated Learning: A Developmental Study from Birth to Senium. By Miriam E. Hebron. London: Methuen, 1966. Pp. 264. £1.75.

Dr Hebron has set herself a formidable task in attempting to give a coherent account of the development of motivation and learning throughout the human life-span. She starts with a brief but lucid account of the physiological basis of arousal, relating this to Eysenck's introversion–extraversion dimension of conditionability. A definition of her subjects is given in terms of preferential modes of processing sensory data. The bulk of the book is devoted to an account of infancy, childhood and adolescence, which leans mainly, although not exclusively, on data from observational and longitudinal studies. The final two chapters are concerned with questions about intellectual and social achievements in maturity and old age. Dr Hebron manages to condense an enormous amount of very disparate information, ranging from reverberating circuits to a schizophrenic baby's body-image and the findings of questionnaire studies of marital adjustment in later life. The result is necessarily uneven if only because of the very different theoretical and empirical bases from which her data are culled. Not surprisingly, too, in an overview of such a vast field, the coverage, detail and recency of information given vary greatly. There is little critical evaluation of the theories and data which are used for expository purposes; and there are serious omissions; for instance, on the central topic of selective attention in learning. This is perhaps inevitable, but it makes the book of doubtful value for senior students of psychology and education to whom it is addressed. However, to have distilled such a variety of material into a connected discourse about development is in itself an achievement.

SUSANNA MILLER

Personality Structure and Measurement. By Hans J. Eysenck and Sybil B. G. Eysenck. London: Routledge & Kegan Paul, 1970. Pp. xiii+365. £3.15.

In 1953, in the first edition of *The Structure of Human Personality*, Professor Eysenck wrote: 'factorial research should be co-operative research . . . is not likely to prosper unless the whole departments, teaching staff and students alike, devote their research energies to the programme research . . . to the unified attack extending over a number of years, on a single important psychological problem'.

Seventeen years and almost as many books later, we can ask how far Professor Eysenck has got. Alas, this volume offers us no more than a fragmentary and uncoordinated glimpse of work in progress. It falls into four quite distinct parts. The first seven chapters are a historical examination of the literature by Professor Eysenck, material which the same author has run through before. The following four chapters are an account also by Professor Eysenck alone of the development and validation of the Maudsley Personality Inventory, together with an epilogue. So far the book has a certain unity, which is faintly echoed in the next three chapters by a group of workers which justify if they do not describe the development of the Eysenck Personality Inventory.

The next five chapters, however, are essentially unrelated factorial studies by a different working group under Professor Eysenck, comparing the Guilford, Cattell and Eysenck Personality Inventories. These, with the technical appendix, could stand alone as an independent work. The last part of the book, ostensibly on 'Personality in Children', consists of one paper by Rachman on neuroticism in children, and two by S. B. G. Eysenck on personality dimensions in children and in subnormal adults. The publication therefore consists of a number of papers by different hands on a wide range of studies, which use factorial methods and personality inventories, but are otherwise independent except in that they form part of a unified attack as Professor Eysenck planned.

From this point of view, the book is disappointing. It suffers, too, from a defect common to any grand design which must take years to carry out. The tools are designed when the job is begun, and though they may be very effective in their day and never lose their edge, they are overtaken by developments in technique. The Atlas computer has been used to reduce the arithmetical labour of handling large numbers of individual scores, but no attempt has been made to use it to carry out more subtle statistical techniques, in spite of a superficial note on the beginnings of the development of numerical taxonomy.

Nevertheless, with care and a knowledge of the background, it is possible to construct a picture of what Professor Eysenck is doing. In the British empirical tradition he has begun to build up a classification of human behaviour which is based not on mythology but on the careful measurement of types of behaviour which the non-specialist and even the layman can recognize with the naked eye and accept as real. More, he has now provided a sensitive and reliable measuring instrument which has shown its validity in practical situations outside its birthplace.

JOHN JAMES

The Psychological World of the Teenager. By Daniel Offer. New York: Basic Books, 1969. Pp. xiv+286. $7.95.

For this reviewer Dr Offer's book has many potential uses; perhaps this fact will colour the review unduly. Unfortunately, the psychological world of the teenager studied by Daniel Offer and colleagues belongs only to boys in the age-group 14 to 18, and almost entirely to the middle class in two separate American suburban areas. Sound reasons for this are stated in the study—reasons which would apply to some areas of this country: most of Offer's adolescent cases referred for psychological problems have come from the middle classes. Some of us are in clinics seeing a fair cross-section socially; some have a vested interest in abnormal girls, and some in the delinquent section drawn disproportionately from the lower strata of the working classes. But this is but another step which Offer predicts—to extend the normative data, and find in the process whether the research techniques are applicable to, or easily modifiable for, a modal population of teenagers, of both sexes, that is by tradition less good at verbalizing, and less geared to self-examination.

Offer and his collaborators screened two large samples of American boys ($n = 326$) at the beginning of their high-school career, by means of a self-image questionnaire. The modal group had to satisfy the statistical requirements on nine out of ten personality scales within the test. The consequent 106 'normals' were then reduced by 3 (only) who were showing considerable behaviour problems in school.

The total sample was followed through at prescribed intervals during 3½ years with a programme of six semi-structured, 45-minute psychiatric interviews, a background information interview, psychological testing, and an interview with the parents. Each student was seen on average three times a year.

The results are fascinating. The students were found to be remarkably homogeneous in behaviour, interests and attitudes, and were closely identified with their parental attitudes. A few acted out their aggressive feelings in delinquent acts and learned from the negative consequences. They complied with the existing social order. 'They did not experiment with drugs.' Only two of the total were actively interested in social-action groups during that phase of their lives. Student–teacher relationships were tense, but in retrospect the lads felt they enjoyed the high-school years. Only 6 per cent had homes broken by death, separation or divorce. A few had traumatic experiences during the research period, but boys and parents coped, and there was generally—as one mother put it—'good, wide-open communication' between the children and parents. They were an action-orientated group, working off aggression and seeking to achieve status, especially in sport. Parents showed some dissatisfaction that their sons did not study harder, and felt life was in some ways too easy for the young. In their sexual behaviour they were 'conservative and restrained'. Eighty per cent approved of premarital sexual intercourse, but only after high school. They stated they were 'not mature enough' for handling the dangers of intercourse.

Results of Thematic Apperception Testing were, like the boys' social attitudes and interests, homogeneous, but the Rorschach test stimulated a wide range of responses. They in fact experienced the classic problems of adolescence, but they coped, each in his own way, yet with certain acceptable techniques in common, such as humour and charm. These they exercised on the interviewing team, whose delight in the research relationship comes across in an infectious way to the reader. By the end of the study the writers can conceptualize that 'adolescence as a period of growth can be undergone without serious disruptions between generations or between the adolescent and his former identity'. 'A state of inner turmoil need not be the password of adolescence'. A comfort to those of us who have had a hunch that this is so for a sizable percentage of young people; and a comfort to us as psychologists that our cases are mainly well outside the modal limits.

HELEN RICHARDSON

Explorations in Communication. Edited by Edmund Carpenter and Marshall McLuhan. London: Cape, 1970. Pp. 208. £1.50.

This book is an anthology of articles selected from the pages of *Explorations*, a journal on communications published at the University of Toronto between 1953 and 1959. A quarter of the papers are concerned with the far-reaching effects new media of communication exert on human living; this of course is the theme McLuhan has expanded in a number of books, with particular reference to the invention of printing and today's electronic media. But many other themes are explored, Buddhist symbolism, space conception, the language of poetry, pure colour, James Joyce and several others. Comprehensive though the list of topics may be, the papers are pervaded by a common spirit of inquiry.

Perhaps the central concept of the book is the 'cultural axiom', a phrase used by Dorothy Lee in one of her admirable contributions. Every culture, including our own, sees reality in terms of a group of axioms of which it is for the most part unconscious and whose validity it assumes to be universal. The projection of such axioms on to a communication conceived in terms of another set is one of the most potent sources of misunderstanding.

In her first paper Mrs Lee draws attention to a basic tenet that underlies the thinking of the Wintu Indians in California. This tenet, that reality exists irrespective of man, sounds simple enough till its full implications are grasped: if reality is indeed wholly independent of man, man's perceptions of reality must be arbitrary and capricious and his media of communication will reflect this arbitrariness. The language of the Wintus is in fact highly subjective, its primary purpose being to convey the speaker's impressions of the moment rather than any information about the external world. In her second paper Mrs Lee is concerned primarily with one of our own axioms, the concept of linearity. Western culture, she says, is dominated by the need to trace linear connexions; this need is manifested in a heavy underlining of causality, whenever we speak of physical happenings, and of purpose and motivation when we are discussing human behaviour. We would no doubt say that this preoccupation is evidence of our interest in objective statement, but whether this is true or not it leads us to make nonsense of the utterances of people such as the Trobriand Islanders whose culture is markedly non-linear.

S. Giedion explores an analogous theme in his study of prehistoric art. Primeval art has no background; it never places objects in an immediate surrounding. All linear directions have equal right, as do all surfaces, regular or irregular. Animals on a cave wall or ceiling that appear to us to be standing on their heads did not appear so to those who depicted them. To them they existed in free space, free from the forces of gravity. It was not till the rise of the Sumerian and Egyptian civilizations that the vertical dimension achieved visual dominance and set the pattern of visual perception for several millennia.

Any reader who has been puzzled by the teenage generation's alleged enthusiasm for McLuhan will find the linkage of ideas set out explicitly in the introduction to this book. Whereas the invention of printing fostered private study by providing every student with his personal texts, electronic media, which make possible the simultaneous sharing of up-to-the-minute information, have an opposite effect. This sharing of information and experience puts a premium on togetherness, creating a new 'tribal juxtaposition' where nobody strives for individual excellence. Mediocrity is cultivated as a means to togetherness, and as a goad to the adult world, which, nourished on ideas of self-discipline and personal striving, is seen as a bulwark of individualism. McLuhan sees no future for literacy or privacy; technology, in his view, is preparing the conditions in which the future will reject the past, or anyway that slice of the past responsible for what we call civilization.

While it is salutary to see this transvaluation of values set out so frankly, this need not stop us questioning the logic of some of the statements or the extent to which these values are held and shared. There is not space in a short notice to explore this exploration as fully as we would like, but we may end by indicating a line of possible counter-inquiry. In his shrewd commentary on Mrs Lee's second paper, Robert Graves draws attention to the difficulty of formulating the differences that separate cultures without exaggerating them. Unless a distinction is over-stated it is unlikely to be grasped at all. Once it is grasped we may well jump to the conclusion that two cultures see the world so differently that there is no possibility of communication between them. But things are rarely as extreme as that. There were individuals who thought and studied before the invention of printing just as there have been millions who shunned reflection in spite of the opportunities for it. Those who study new trends often overlook the cybernetic principle which ensures eventual reaction to the trendiest developments. The sophisticated publicist can hardly be unaware of these oscillations; it is his fate to believe fervently in the importance of his message and at the same moment to take it with a pinch of salt.

JOHN PARRY

The Basic Fault. By Michael Balint. London: Tavistock Publications. 1968. Pp. 205. £1.90.

This author is most widely known for the expert help he has given to the non-specialist in psychotherapy, although his earlier works were primarily for the practising analyst. In the present volume he has made another contribution to theory which should be of immense interest, not only to the specialist but to anyone with a reasonable knowledge of the basic psychoanalytic literature. He deals with the phase which may occur in a significant group of patients during the working through of the transference, and in which the patient becomes extremely regressed. Frequently this has led either to breakdown of treatment or to an interminable analysis with great frustration on both sides. Verbal communication may break down and therapist and patient lose contact. Early disasters of this kind led analysts to avoid this state in a patient, so that they were then denied its therapeutic potential. Dr Balint shows convincingly how it may be understood and worked through and also gives, for the first time to my knowledge, an analysis of types of regression. Some of his comments on techniques have affinities with the work of writers who have worked with patients whose illness renders them unable to communicate except on a 'primitive' level, e.g. Leonard B. Hill and Trygve Braatoy. It is this level, which is essentially undifferentiated and preverbal, that is delineated and named 'the basic fault'. Theoretically it has much in common with the schizoid position of Melanie Klein. However, Dr Balint's elucidation has a potential appeal to therapists of all schools and this in itself is a not inconsiderable virtue in a book already valuable for its practical application.

P. M. SHAW

CHANGE-OVER TO SI UNITS

From January 1972 the journals of the British Psychological Society will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Further information about SI Units is contained in the revised *Suggestions to Authors* pamphlet issued by the Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

Authors who refer to physical measurements in their papers should now use SI Units except where it would be absurd to do so. Conversion tables will be published during the change-over period but in *Br. J. soc. clin. Psychol.* these will be restricted to units of length as others rarely occur. When they do, conversion factors will be stated in footnotes. Common units of time (e.g. hour, year) will persist.

1 metre (m) = 39.37 inches
1 metre (m) = 3.28 feet
0.0254m = 1 inch
0.3048m = 1 foot

Developmental Sciences Series

Language Acquisition: Models and Methods

edited by Renira Huxley and Elisabeth Ingram

September 1971, xiv+312 pp., £5.00

The most effective treatment of the subject to date, this volume presents ideas, theories and findings concerning the development of linguistic, cognitive and communicative abilities in the child. It represents one of the few attempts to bring together experts from different disciplines, and reports on the opinions of, and interaction between, these disciplines.

Contents
Sociolinguistics. Grammar. Cognition and language. Clinical. Bibliography. Author index. Subject index.

The Origins of Human Social Relations

edited by H. R. Schaffer

September 1971, xiv+298 pp., £5.00

It has long been believed that the nature of the child's primary social relationship plays a vital role in personality development. How this relationship is established in infancy and the processes responsible for its formation and development are topics that have aroused speculation and controversy and have only recently been subjected to empirical enquiry. The papers contained on this volume represent current attempts to examine these topics and to provide findings that will enable the relationship to be defined more precisely.

Contents
Aspects of early social behaviour. Comparative perspectives. Social perspectives. Basic processes. Author index. Subject index.

Academic Press

London and New York
Berkeley Square House, Berkeley
Square, London W1X 6BA, England
111 Fifth Avenue, New York
N.Y. 10003, USA

The Journal of SOCIAL ISSUES

WINTER 1971 VOL. XXVII NO. 1

The New Left and The Old
Issue Editor: Armand L. Mauss

The Lost Promise of Reconciliation: New Left vs. Old Left . . . . . Armand L. Mauss
The New Left and American Politics after Ten Years . . . . R. Flacks
Political Libertarianism at Berkeley: An Application of Perspectives from the New Student Left . . . . Henry C. Finney
Criteria and Contingencies of Success in a Radical Political Movement . . . N. J. Demerath III, Gerald Marwell, and Michael T. Aiken
Majority Involvement in Minority Movements: Civil Right, Abolition, Untouchability . . . . Gary T. Marx and Michael Useem
Old and New Left Activity in the Legal Order: An Interpretation . . . . Nathan Hakman
Romantics and Renegades: Political Defection and the Radical Left . . . Gilbert Abcarian
New Left as Right: Convergent Themes of Political Discontent David R. Schweitzer and James M. Elden
The Student Left in Puerto Rico . . . . Arthur Liebman
On Being Strangled by the Stars and Stripes . . . . Armand L. Mauss
Biographical Sketches
Comments and Rejoinders
Fred Arnstein . . . *Comment on* "Three Not-So-Obvious Contributions of Psychology to Peace" by Ralph K. White
Ralph K. White . . . *Rejoinder*
Fred Arnstein . . . *Response to White's Rejoinder*
Ralph K. White . . . *Counter Response*
H. Edward Ransford . . . *Comment on* "Internal-External Control and Black Militancy" by John R. Forward and Jay R. Williams
John R. Forward and Jay R. Williams . . . *Rejoinder*

Order from: Journal of Social Issues Post Office Box 1248 Ann Arbor, Michigan 48106
1971 *Annual Subscriptions Rates* INSTITUTIONS $15.00 INDIVIDUALS $9.00
SINGLE COPIES $3.00

Published quarterly by
The Society for the Psychological Study of Social Issues
(a division of the American Psychological Association)

JAPANESE PSYCHOLOGICAL RESEARCH

Vol. 12, no. 4 (Dec., 1970)

Yanai, H. Factor Analysis with External Criteria: Application of Analysis of Variance and Regression Analysis Technique to Factor Analysis

Takahashi, S. Analysis of Weighted averaging Model on Integration of Informations in Personality Impression Formation

Oyama, T. The Visually Perceived Velocity as a Function of Aperture Size, Strip Size, Luminance, and Motion Direction

Tanaka, Y., and Nakajima, Y. Perception of Moving Figures 1; On the Reduction Phenomena

Iwawaki, S., Oyama, T., Sugiyama, Y., Kikuchi, M., and Komatsu, R. Development and Validation of the Japanese Version of the MPI (Maudsley Personality Inventory)

Indow, T. The Formation of Subjective Value: A Theoretical and Empirical Analysis

Subscription price is $7.00 per year. We offer a special rate of $6.00 for members of IUPS member societies. Single copy price is $2.00.

Send Subscription Orders to:

JAPANESE PSYCHOLOGICAL ASSOCIATION
37-13-802, Hongo 4 chome, Bunkyo-ku, Tokyo 113, Japan

Personality: An International Journal

Vol. 2, No. 2, Summer 1971

CONTENTS

Similarity-Attraction or Demand Characteristics? 77

John Lamberth and Donn Byrne

Physiological Response as a Function of Psychosexual Fixation 93

Jerry Higgins, David M. Kuhlman, and Marvin B. Eisen

A Factor Analytic Study of Interests 97

C. J. Adcock and G. C. Townsley

An Item-Factor-Analysis of the 16 PF 117

Edgar Howarth and James A. Browne

Personality Correlates of Behavior Problems in Early Childhood 141

T. E. Dielman, R. B. Cattell, and Carolyn Lepper

The Dimensions of Pathology: Proof of Their Projection Beyond the Normal 16 PF Source Traits 149

Karl H. Delhees and Raymond B. Cattell

Interpersonal Attraction Toward a Black Person as a Function of Value Importance 175

Louis A. Penner

Published quarterly by Plenum Publishing Corporation, 227 West 17th Street, New York, N.Y. 10011 under the auspices of the Society for Personality Research Incorporated.

Subscription price: Volume 2, 1971 (four issues) $12.00.

Br. J. soc. clin. Psychol. (1971), **10**, pp. 201–212
With 1 *text-figure*
Printed in Great Britain

Social Attitudes and Social Class

BY H. J. EYSENCK

University of London

In the course of an extensive survey, made by the Audience Research Department of the B.B.C., some 2000 male and female subjects, the members of the B.B.C. Viewing Panel, were administered an attitude questionnaire containing 28 questions. Subjects were divided into three social-class groups (middle class, skilled working class, unskilled working class), three age-groups and two sex-groups; answers to the questionnaire items were correlated and factor-analysed for the resulting 18 groups separately. Sex and age did not seem to have any great effect on the patterns of intercorrelations emerging, and consequently three samples were constructed of subjects in the three social-class groups in such a way that proportions of men and women, and of the three age-groups, were equal. These three groups were then submitted to a factorial analysis, and factor similarity indices were calculated between the two higher-order factors which emerged for each group to indicate to what extent attitude structure might differ from class to class; no systematic differences were observed, other than a tendency for attitudes to be more strongly structured in the higher social-class groups. In view of the great similarities between the three solutions for the class groups separately, they were combined into one group, and a factor analysis carried out on the resulting correlations. The structure observed was found to be very similar to that reported earlier in *The Psychology of Politics*, although an alternative descriptive system to the R and T factor system is also discussed. Age was found to correlate with conservatism, male sex with toughmindedness. The results are interpreted as suggesting that there have been no systematic changes in the structure of social attitudes in this country in the 20 years that have elapsed since the research published in *The Psychology of Politics* was carried out.

The present study is an extension of work reported in *The Psychology of Politics* (Eysenck, 1954) on the organization of social attitudes; particular attention is paid to the problems raised in this connexion by social class differences (Eysenck, 1951). Essentially what the writer discovered in these studies was (*a*) that when relations between large samples of questions concerning social issues are studied and factor-analysed, two main factors emerge which are orthogonal to each other; these have been denoted R (radicalism–conservatism) and T (toughmindedness–tendermindedness). In addition it was found (*b*) that when equated for political voting pattern, working-class subjects were more conservative and more toughminded than middle-class subjects; this was equally true of conservative, liberal, labour and communist voters. These studies left several questions unanswered, and it is the purpose of this paper to raise some of these and attempt to answer them by reference to an empirical investigation.

The first question to arise refers to the relation between the two-factor pattern and the social class of respondents. Factor analyses have hitherto been carried out largely on student groups or on mixed social groups; it does not follow that identical structures of attitudes will be found in middle-class groups, skilled working-class groups, and unskilled working-class groups. It is possible that the structures found in the past are typical of middle-class samples, or are an artifact of throwing together subjects from different social classes each having a different attitude structure; there is no empirical evidence to throw any light on this question, and it seemed worth while to carry out an experiment to test the hypothesis that social classes do not differ in respect of attitude organization. Even if the general structure of attitudes were found to be the same, it might still be possible that this organization was tighter among middle-class subjects than among working-class subjects; that this might be so is suggested by the well-known fact that working-class subjects tend to give more 'don't know' answers in opinion surveys.

The second question to arise relates to a widespread belief that working-class subjects are more 'left-wing' than middle-class subjects, and consequently more 'progressive'; the evidence certainly suggests that they vote Labour in greater proportion than do middle-class subjects, and the Labour Party, almost by definition, is less conservative than the Conservative Party (Eysenck, 1947). Yet public opinion poll results often show working-class subjects to be more ethnocentric and less 'advanced' in respect of such issues as the death penalty, the treatment of conscientious objectors, or sexual morals than middle-class subjects (Eysenck, 1970); the fact that the writer found them to be more conservative in all matters other than economic ones has already been mentioned. It seemed worth while to investigate this problem in some more detail, using random samples of working-class and middle-class subjects. This seemed all the more necessary as some 20 years had elapsed since the writer's last investigation, and many changes could have taken place since then; it cannot be assumed that findings in social psychology remain unaffected and unaltered by events taking place in time.

METHOD

A set of 28 questions taken from or adapted from the questionnaire used in *The Psychology of Politics* was used for this investigation; the questions used are given in full in Table 1.

Table 1

1. People should realize that their greatest obligation is to their family.
2. Production and trade should be free from government interference.
3. An occupation by a foreign power is better than war.
4. Men and women have the right to find out whether they are sexually suited before marriage (e.g. by trial marriage).
5. Nowadays, more and more people are prying into matters which do not concern them.
6. Jews are as valuable citizens as any other group.
7. The death penalty is barbaric, and it was right to abolish it.
8. Most religious people are hypocrites.
9. Our treatment of criminals is too harsh, we should try to cure them, not punish them.
10. Sex relations except in marriage are always wrong.
11. It would be best to keep coloured people in their own districts and schools, in order to prevent too much contact with whites.
12. Compulsory military training in peace-time is essential for the survival of this country.
13. Sex crimes, such as rape and attacks on children, deserve more than mere imprisonment, such criminals ought to be flogged, or worse.
14. Persons with serious hereditary defects and diseases should be compulsorily sterilized.
15. It would be a mistake to have coloured people as foremen over whites.
16. 'My country, right or wrong' is a saying which expresses a fundamentally desirable attitude.

Table 1—*continued*

17. The idea of God is an invention of the human mind.
18. A person should be free to take his own life, if he wishes to do so, without any interference from society.
19. Free love between men and women should be encouraged as a means towards mental and physical health.
20. A white lie is often a good thing.
21. The so-called underdog deserves little sympathy or help from successful people.
22. The Church should attempt to increase its influence on the life of the nation.
23. Coloured people are innately inferior to white people.
24. The dropping of the first atom bomb on a Japanese city, killing thousands of innocent women and children, was morally wrong and incompatible with our kind of civilization.
25. All forms of discrimination against the coloured races, the Jews, etc., should be made illegal, and subject to heavy penalties.
26. Capitalism is immoral because it exploits the worker by failing to give him full value for his productive labour.
27. Christ was divine, wholly or partly in a sense different from other men.
28. The maintenance of internal order within the nation is more important than ensuring that there is complete freedom for all.

This inventory, together with certain other questions, was administered to a random sample of the adult population; also ascertained were their age, sex and social class. Age-groups were subdivided for the purpose of this study into young (21–25), middle-aged (26–45) and old (46–65); the terms and divisions are of course arbitrary. Social classes were also divided into three: middle class, including higher professional and administrative workers, lower professional, technical and executive workers; skilled workers, including workers from highly skilled to moderately skilled; and unskilled workers, including semi-skilled workers. These 2 × 3 × 3 groups contained unequal numbers of subjects; the actual numbers are given in Table 2.

Table 2. *Male and female groups, numbered from 1 to 9, showing numbers of subjects in each group*

	Male			Female		
	Young	Middle-aged	Old	Young	Middle-aged	Old
Middle class	37 (1)	148 (2)	106 (3)	13 (1)	93 (2)	41 (3)
Skilled	116 (4)	336 (5)	308 (6)	158 (4)	391 (5)	240 (6)
Unskilled	108 (7)	127 (8)	106 (9)	128 (7)	240 (8)	206 (9)

RESULTS

Replies were coded from 1, indicating strong agreement, through 2, signifying agreement, 3, signifying neither agreement nor dissent, and 4, signifying disagreement, to 5, signifying strong disagreement. Means for the nine male groups are given in Table 3, and for the nine female groups in Table 4. The results show clearly that classes differ from each other, that age-groups differ, and also that men and women differ in their endorsements. The tables speak for themselves, but a brief discussion may be helpful; in drawing attention to the most marked differences we have in each case tried to make comparisons which would be valid with the other variables held constant. In other words, differences between men and women are commented upon when they are found across class and age-groups, and so on for the other differences in question.

Table 3. *Mean scores on 28 items of nine male groups*

	Group								
Item	1	2	3	4	5	6	7	8	9
1	2·24	1·89	1·88	1·80	1·64	1·54	1·83	1·49	1·52
2	3·02	2·77	2·54	2·42	2·48	2·32	2·63	2·38	2·54
3	3·76	3·94	3·89	3·94	4·07	4·04	3·81	4·06	3·96
4	2·54	3·01	3·35	2·70	3·18	3·76	2·72	3·28	3·52
5	2·54	2·49	2·47	2·37	2·23	2·21	2·08	1·99	2·15
6	1·65	1·69	1·87	1·94	1·82	1·79	1·95	1·84	1·80
7	2·81	3·34	3·42	3·55	3·73	3·70	3·76	3·84	3·87
8	3·14	3·53	3·65	3·13	3·25	3·33	3·05	2·95	3·18
9	3·11	3·64	3·68	3·73	3·97	3·90	3·73	3·87	3·96
10	4·14	3·36	3·22	3·74	3·47	2·81	3·69	3·35	3·05
11	4·38	4·01	3·67	3·79	3·63	3·50	3·58	3·60	3·40
12	3·76	3·41	3·05	3·60	2·78	2·59	2·96	2·57	2·25
13	3·41	2·73	2·75	2·72	2·07	1·98	2·08	1·78	1·76
14	3·30	3·11	2·79	3·41	2·82	2·50	2·94	2·85	2·42
15	3·59	3·77	3·42	3·37	3·32	2·95	3·06	3·08	2·87
16	3·54	3·51	3·43	3·34	3·24	3·18	2·81	3·06	3·00
17	2·59	3·20	3·28	3·14	3·20	3·48	3·27	3·15	3·38
18	2·78	3·05	3·12	2·96	3·17	3·24	3·06	3·09	3·19
19	3·73	3·91	3·96	3·41	3·77	4·06	3·31	3·58	3·63
20	2·11	2·24	2·21	2·15	2·19	2·26	2·35	2·14	2·05
21	3·86	3·90	3·80	3·67	3·70	3·84	3·66	3·57	3·43
22	3·30	2·78	2·37	3·44	3·09	2·58	3·22	3·08	2·81
23	4·51	4·15	3·92	3·91	3·87	3·61	3·75	3·87	3·79
24	2·81	3·19	2·88	2·98	3·19	2·90	2·84	2·88	2·56
25	2·70	2·86	2·85	2·83	2·86	2·70	2·79	2·66	2·53
26	3·62	3·90	3·84	3·18	3·53	3·56	2·59	3·00	2·90
27	2·89	2·64	2·57	2·92	2·76	2·44	2·74	2·76	2·53
28	2·89	2·41	2·25	2·91	2·69	2·28	2·94	2·79	2·23

Table 4. *Mean scores on 28 items of nine female groups*

	Group								
Item	1	2	3	4	5	6	7	8	9
1	2·15	1·90	1·78	1·87	1·72	1·79	1·77	1·53	1·62
2	3·00	2·70	2·29	2·53	2·40	2·18	2·59	2·41	2·27
3	3·92	3·69	3·51	3·78	3·75	3·76	3·41	3·77	3·87
4	2·46	3·54	3·63	3·06	3·57	3·90	3·17	3·60	4·02
5	3·08	2·66	2·15	2·18	2·43	2·38	2·16	2·19	2·17
6	1·69	1·55	1·73	1·80	1·75	1·79	2·01	1·90	1·93
7	2·62	2·76	3·15	3·82	3·49	3·52	3·87	3·74	3·79
8	3·85	3·84	4·00	3·27	3·51	3·72	3·08	3·18	3·52
9	2·92	3·22	3·51	3·92	3·83	3·82	3·92	3·92	3·84
10	3·77	3·20	3·10	3·59	3·03	2·64	3·40	2·89	2·89
11	4·62	4·16	4·02	3·89	3·71	3·45	3·67	3·50	3·31
12	3·08	3·19	3·17	2·64	2·74	2·70	2·56	2·51	2·54
13	3·46	2·98	2·88	2·13	2·17	2·05	1·84	1·72	1·76
14	3·62	3·09	2·71	3·00	2·83	2·43	2·89	2·59	2·45
15	4·15	3·80	3·37	3·37	3·42	3·00	3·27	3·09	2·78
16	4·00	3·46	3·59	3·06	3·21	3·09	3·04	2·94	2·94
17	3·77	3·51	3·73	3·53	3·71	3·80	3·85	3·58	3·73
18	3·15	3·19	2·93	3·01	3·10	3·10	3·01	3·10	3·22
19	3·77	4·20	4·34	3·93	4·10	4·17	3·73	4·02	4·11

Table 4—*continued*

Item	Group 1	2	3	4	5	6	7	8	9
20	2·08	2·16	2·10	2·15	2·12	2·17	2·24	2·17	2·17
21	4·23	3·77	4·00	3·57	3·60	3·72	3·24	3·46	3·56
22	2·77	2·59	2·12	3·22	2·87	2·46	2·96	3·05	2·62
23	4·62	4·46	4·17	4·16	4·06	3·94	3·79	3·80	3·86
24	1·92	2·42	2·51	2·31	2·56	2·51	2·34	2·38	2·35
25	2·31	2·60	2·54	2·69	2·73	2·51	2·77	2·76	2·61
26	3·69	3·70	3·68	3·04	3·48	3·46	2·82	3·13	3·14
27	2·23	2·31	2·37	2·51	2·40	2·10	2·51	2·48	2·25
28	2·69	2·71	2·32	2·98	2·87	2·51	3·12	2·87	2·57

DISCUSSION

Comparatively speaking then, middle-class people favour trial marriages, do not think that people are prying into their affairs, believe that Jews are valuable citizens, that the death penalty is barbaric, and that our treatment of criminals is too harsh. They do not think that religious people are hypocrites, or that extra-marital sex is wrong; they do not want to keep coloured people separate, do not favour compulsory military training, flogging or compulsory sterilization. They are not opposed to coloured foremen, and do not believe in the saying 'My country, right or wrong'. They believe that underdogs deserve sympathy, that the Church should increase its influence, but not that coloured people are inferior, or that capitalism is immoral; they do consider that order is more important than complete freedom. Working-class people hold the opposite views on these matters; in other words, they emerge on most issues as conservative and toughminded, very much as in the writer's original studies. Thus working-class people, in summary, are nationalistic, even jingoistic, xenophobic, antisemitic, racialist, inhumane, narrowly moralistic in sexual matters, and unconcerned with ethical or religious issues.

Sex differences are rather as one might have expected, with women emerging as tenderminded conservatives. Women think that occupation is better than war, that the death penalty is barbaric, but not that religious people are hypocrites. They consider extra-marital sex wrong, do not agree that God is an invention of the human mind, or that free love is a good thing. They would like to see the Church increase its influence, and consider the dropping of the atom bomb immoral; they are against racial discrimination, and do not consider coloured people inferior; they also consider Christ divine. Age, on the other hand, seems to make people both more conservative and more toughminded. Older respondents believe in one's obligations to the family, resent government interference, do not believe in trial marriages, do not consider religious people hypocrites, think that extra-marital sex is wrong, want to keep coloured people separate, believe in compulsory sterilization, flogging, and the inferiority of coloured people; they also do not approve of coloured foremen. They do not hold with free love, want the Church to increase its influence, consider order more important than freedom, and do not feel that capitalism is immoral. This study is of course not a longitudinal one; we cannot say whether age *per se* produces such changes in attitudes, or whether the older people grew up in a different climate of opinion and simply maintained these attitudes over the years. Perhaps both influences were at work in producing the observed effect.*

We must now turn to a consideration of the intercorrelations between items, and the factor analysis of the resulting matrices of correlations. In the first place, such analyses, using principal components methods of extraction and rotation by Promax, were performed on all 18 groups independently; some of these are of course too small to make it possible to take the results very seriously, but for most groups numbers are sufficient for interpretation. It would be quite impossible to print all the results of these analyses, or even to discuss them in detail; suffice it to say that in all groups two latent roots stood out beyond all the rest as being consistently much larger, and that for all groups there was considerable similarity in factor composition for the two factors corresponding to these two roots, when rotated by Promax. This suggests that similar principles of organization hold across sex, age and class groupings. A more detailed presentation of results affecting class will be given presently.

In order to reduce the vast amount of data, it was decided to carry out a new set of factor analyses on the three social-class groups, collapsed over sex and age groupings. Proportional samples were taken on a random basis from the three class groups, resulting in 382 middle-class subjects, 1235 skilled subjects, and 553 unskilled subjects. (Proportional samples had to be taken for each group in order to have the same proportions of men and women, old and young, in each of the samples, as otherwise differences between class groups might have been caused by their differential make-up along the sex or age dimension.) Including age and sex in the analysis, there were 30 variables, and consequently 30 factors were extracted from each matrix of intercorrelations by principal components analysis. Nine or ten of these factors, having latent roots exceeding unity, were then rotated by Promax, and higher-order factors extracted. The same process was followed with a combined set of middle-class, skilled and unskilled working-class group, made up of the three separate groups; this will be referred to as the combined group. In view of the similarity of the primary factors extracted from the different groups, we will only discuss those extracted from the combined group; higher-order factors will be given for all groups.

Table 5 gives the latent roots for the first 10 factors extracted in order, for the combined group and also for the three social-class groups. Two facts stand out from this table; there are nine or 10 factors in each sample which have latent roots exceeding unity, so that something like eight factors might be regarded as significant, and the two first latent roots are far and away the biggest. It is of course a necessary feature of the method of extraction that the first two latent roots should be the biggest, but attention is here called to the difference in size between them and the subsequent ones; there is an unusually precipitate drop from root 2 to root 3, suggesting that two roots mediate the main content of this particular universe. In line with this fact is the discovery that the higher-order analysis also leads to two factors for the three separate class groups, and for the combined groups.

---

* Attitudes characteristic of the working class, the old and male respondents can be seen to favour authoritarian views; together they provide a picture of the old, male working-class person which resembles Alf Garnett too closely for comfort.

What are these eight primary factors? Factor I resembles the well-known F or authoritarian factor; items having loadings on this are: People should realize that their greatest obligation is to their family (0·56), production and trade should be free from government interference (0·69), more and more people are prying into matters which do not concern them (0·56), compulsory military training is necessary (0·50), sex criminals should be flogged (0·52), and 'my country, right or wrong' (0·49). Factor II is a religious one, loading on the following: Religious people are not hypocrites (0·64), the idea of God is not an invention of the human mind (0·81), the Church should increase its influence (0·70), and Christ was divine (0·83). Factor III is one of ethnocentrism: Jews are not valuable citizens (0·65), coloured people should be kept in their own districts (0·70), coloureds should not be foremen over whites (0·64), coloureds are innately inferior (0·68), and discrimination should not be penalized (0·63). Factor IV is one of humanitarianism: Capitalism is immoral (0·80), dropping the atom bomb was morally wrong (0·37), occupation by a foreign power is better than war (0·43), and production and trade should not be free from government interference (0·46). This factor is only named with diffidence, as only one loading is at all large. Factor V is one of sexual morals: Trial marriage (0·81), sex relations outside marriage not wrong (0·76), and free love should be encouraged (0·78). Factor VI is one of toughmindedness: Death penalty is not barbaric (0·70), treatment of criminals is not too harsh (0·77), sex criminals should be flogged (0·49), and atom bomb dropping was not immoral (0·40). Factors VII and VIII have their main loadings on age and sex, respectively, and are therefore of comparatively little interest; sex and age differences have already been discussed.

Table 5. *Latent roots of first 10 factors extracted for middle-class, skilled and unskilled groups, as well as for the combined group*

	Latent roots			
	Total group	Middle class	Skilled	Unskilled
1	4·2	4·9	4·1	3·4
2	3·5	3·5	3·6	3·1
3	1·7	1·8	1·7	1·9
4	1·5	1·6	1·5	1·6
5	1·4	1·4	1·4	1·4
6	1·3	1·3	1·3	1·3
7	1·2	1·1	1·1	1·2
8	1·1	1·1	1·1	1·2
9	1·0	1·0	1·0	1·2
10	0·9	1·0	1·0	1·0

These factors are of course not orthogonal; Promax rotates factors into optimum simple structure regardless of orthogonality. The correlations observed would seem to make good sense. Authoritarianism (F) correlates 0·32 with ethnocentrism; religious attitudes correlate −0·43 with sexually permissive ones. Most of the other

Table 6. *Factor loadings in Factors I and II of combined groups* (1 *and* 5) *and of middle-class* (2 *and* 6), *skilled working-class* (3 *and* 7) *and unskilled working-class* (4 *and* 8) *subjects*

		Factor I				Factor II		
	1	2	3	4	5	6	7	8
Age	−0·22	−0·22	−0·27	−0·02	0·41	0·24	0·44	0·50
Sex	0·38	0·28	0·32	0·01	0·36	0·40	0·26	−0·02
1	0·33	0·27	0·29	0·21	−0·18	−0·16	−0·17	−0·31
2	0·34	0·39	0·44	0·42	−0·20	−0·24	−0·15	−0·24
3	−0·42	−0·26	−0·38	0·19	−0·07	0·08	−0·09	−0·15
4	−0·12	−0·18	−0·08	−0·02	0·69	0·64	0·73	0·50
5	0·29	0·31	0·30	0·23	−0·02	−0·11	0·05	0·07
6	−0·23	−0·38	−0·17	−0·48	−0·13	−0·09	−0·10	−0·22
7	−0·65	−0·62	−0·59	−0·33	−0·04	−0·09	−0·05	−0·08
8	0·19	0·13	0·17	0·10	0·48	0·53	0·49	0·40
9	−0·58	−0·64	−0·50	−0·22	−0·00	0·06	0·00	−0·02
10	0·09	0·14	0·04	0·13	−0·68	−0·62	−0·70	−0·59
11	0·79	0·58	0·53	0·71	−0·06	−0·03	−0·05	−0·10
12	0·38	0·20	0·38	0·13	−0·18	−0·35	−0·13	−0·22
13	0·57	0·49	0·53	0·41	−0·09	−0·18	−0·08	−0·13
14	0·29	0·31	0·35	0·32	−0·02	0·10	0·03	−0·30
15	0·49	0·54	0·49	0·67	−0·09	0·01	−0·11	−0·09
16	0·35	0·42	0·31	0·21	−0·08	−0·09	−0·07	−0·04
17	−0·01	0·00	0·00	0·18	0·66	0·73	0·65	0·40
18	0·00	0·10	0·10	0·24	0·35	0·50	0·45	0·06
19	−0·01	−0·09	−0·04	0·20	0·50	0·48	0·49	0·26
20	0·03	0·21	0·03	−0·12	0·12	0·20	0·24	0·18
21	0·14	0·42	0·17	0·31	0·06	0·08	0·09	0·17
22	−0·02	0·02	0·02	−0·09	−0·60	−0·58	−0·61	−0·61
23	0·50	0·56	0·52	0·55	0·08	0·08	0·09	0·08
24	−0·54	−0·54	−0·54	−0·13	−0·25	−0·23	−0·24	−0·30
25	−0·40	−0·47	−0·48	−0·36	−0·23	−0·33	−0·23	−0·35
26	−0·18	−0·26	−0·39	−0·13	0·12	0·01	0·09	0·02
27	−0·00	−0·05	0·00	−0·13	−0·65	−0·69	−0·66	−0·33
28	0·33	0·35	0·34	−0·04	−0·08	−0·03	−0·10	−0·25

correlations are rather small, but the matrices of intercorrelations can be factor-analysed to reveal higher-order factors. Table 6 gives loadings on the two higher-order factors which emerge from such an analysis; first- and second-factor loadings are given for the combined groups, and also for the three separate social-class groups. These two factors are almost entirely orthogonal in each case; correlations between them are are −0·07 for the combined groups, and −0·18, 0·05 and 0·02 for the three social-class groups. Are the factors extracted from the different classes identical? Indices of factor comparison (Kaiser *et al.*, 1969) were calculated; they amounted

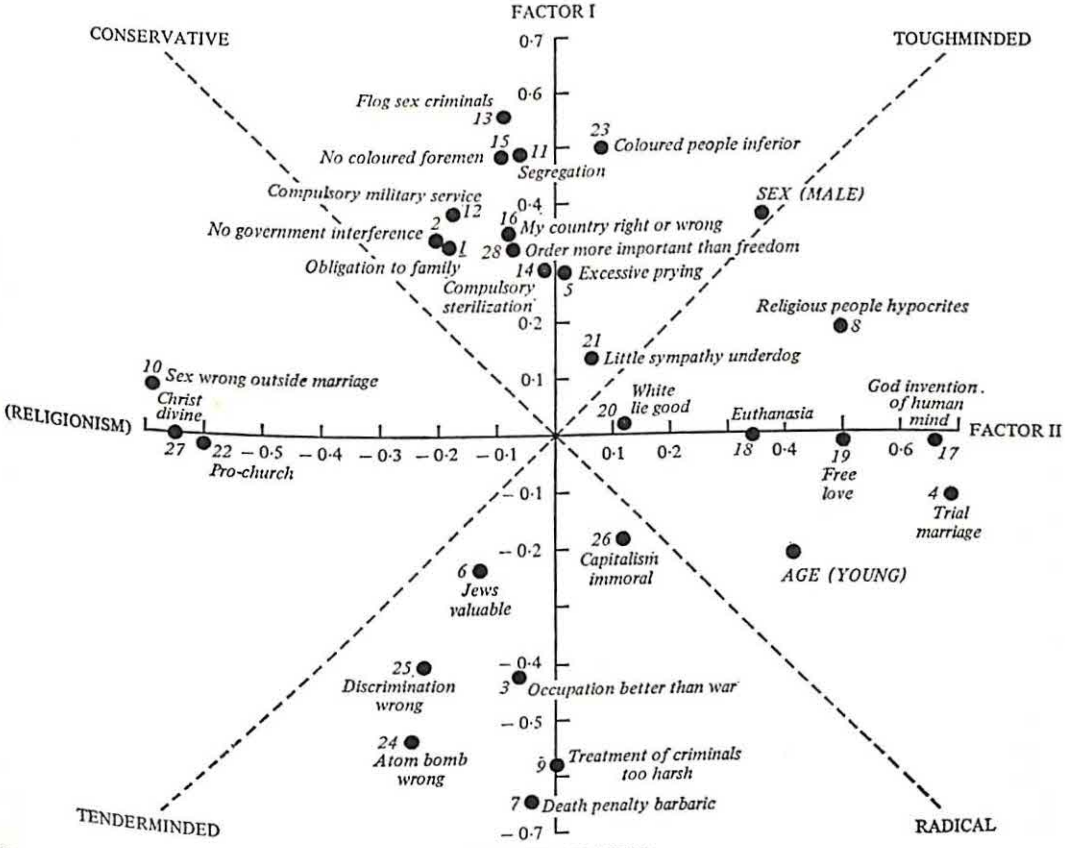


Fig. 1. Diagram representing two main factors underlying the organization of social attitudes.

to 0·99 or better in each case, demonstrating that different social classes share identical attitude structure, at least as far as the attitudes sampled here are concerned. This answers our main question; analysis of social attitude structure reveals two main, independent factors in each case, and these factors are not artifacts produced by the throwing together of heterogeneous social groups. We must now turn to the nature of these two factors. In view of the great similarities between the different analyses, only the results for the combined group have been plotted (Fig. 1).

When the configuration of items in Fig. 1 is compared with that shown on page 130 of *The Psychology of Politics*, it will be seen that the two are essentially identical; changes occurring in political and social attitudes during the past 20 years, and differences in the composition of the samples, have not affected the fundamental two-factor structure of this particular universe of questions. There is, however, an interesting shift in the position of the two factor axes; it appears that these have rotated through an angle of approximately 45°, from the position indicated by the broken lines to that occupied by the solid lines. Reasons for this shift will be discussed presently; first let us consider the nature of the factors as they now appear. Factor I opposes authoritarianism and humanitarianism. The authoritarian wants to flog sex criminals, believes coloured people to be innately inferior, favours compulsory military service, dislikes government interference, values order more than freedom, favours segregation of coloureds, and stresses one's obligation to one's family; all this agrees well with the classical 'authoritarianism' picture. The humanitarian believes discrimination on the grounds of colour to be wrong, considers occupation better than war, favours more lenient treatment of criminals, believes the death penalty to be barbaric, and thinks that it was morally wrong to drop the atom bomb. This factor resembles Ferguson's (1939) original suggestion of a 'humanitarian' factor, a discovery which preceded the work on the authoritarian personality but never received the recognition which it deserved.

The second factor resembles Ferguson's other suggestion, namely that of a factor of 'religionism'; this is opposed to attitudes favouring sexual permissiveness. Agreement with items suggesting that Christ was divine, and that the Church should increase its influence, go with belief that extra-marital sex is wrong; agreement with items suggesting that God is an invention of the human mind, and that religious people are hypocrites, goes with belief in trial marriage, free love, and euthanasia. These two factors emerged in Eysenck's original work as the four quadrants of the radicalism–conservatism and toughminded–tenderminded framework; religious beliefs appeared in the tenderminded conservative quadrant, authoritarian beliefs in the toughminded conservative quadrant, humanitarian beliefs in the tenderminded radical quadrant, and permissive beliefs in the toughminded radical quadrant.

Why has there been such a rotation of axes, and which position is right? It is easier to answer the former question than the latter. Simple structure seeks out clusters of items; in the original analysis the conservative–radical factor was centred on questions relating to economic radicalism and conservatism, e.g. 'abolish private property' and 'nationalization is inefficient'. Such questions were omitted in the B.B.C. study, and as a consequence the centre of the R factor cluster is missing; as simple structure seeks out what now appear the main clusters, i.e. the religious *v.* permissive and the humanitarian *v.* authoritarian, such an omission must inevitably lead to a breaking up of the single R factor into two; as shown in *The Psychology of Politics*, the T factor, for reasons discussed there, has no such centre in any case.

Which solution is the better? Statistically there is no meaningful answer to this question: statistics ends with the delineation of the two-dimensional pattern of item locations graphed in Fig. 1, where the scalar products for each pair of items approximate closely to the observed intercorrrelations. A choice has to be made on psychological grounds; this point has been argued already in *The Psychology of Politics*. Thurstone would dispute this suggestion; he believes that the main advantage of simple structure over other methods of rotation lies in the introduction of a statistical principle which eschews subjectivity and is relatively independent of the precise choice of items (or tests) correlated. Unfortunately this is not necessarily so, as the present data show only too clearly; the omission of just a few items may change the simple structure solution in a very clear-cut and predictable manner. Given that psychological considerations are relevant and important, what are the arguments on the two sides? The Ferguson solution has the advantage that the two factors which emerge are clear-cut, orthogonal and easy to understand; they agree with common sense in the way in which they structure the universe of attitudes. In addition they agree with the large body of work which has been done on the authoritarian dimension, thus bringing this into line with factorial attitude research. These are very real advantages, but in the writer's view they are not sufficient to overcome the advantages which accrue to the alternative solution in terms of the R and T factors.

Radicalism and conservatism are concepts which have been used for centuries to organize and conceptualize groups of attitudes; it seems unlikely that political parties would have grown up around concepts which were just combinations of authoritarian religionism (conservatism) or permissive humanitarianism (radicalism). On a Marxist view, and equally on a capitalist view, the concepts of economic radicalism and economic conservatism, associated as they are with socialism and capitalism, are quite central in this field, and a factorial solution arrived at by leaving them out completely from the universe of attitudes sampled is clearly unsatisfactory. The relationship of the R factor to voting patterns has been demonstrated in *The Psychology of Politics*; it would seem unreasonable to give up the advantages which this close and clear-cut relation bestows upon our solution. The same may be said of the factors of age and sex; it seems reasonable to find that men are toughminded, and that the young are radical. Reinterpret these findings in terms of religionism and authoritarianism, and the relationships become much more complex. To the writer, these relations between factors outside influence have always been an important part of the interpretation of factors, and it would appear that only in this way can we escape from the circular argument which factorial analysis constitutes when presented in isolation. This argument may also be extended to personality; it was argued in *The Psychology of Politics* that while the R factor was a genuine social attitude factor, T was the projection on to this field of extraverted and introverted personality patterns—extraverts tending towards tough, introverts towards tender expressions of radical or conservative attitudes. There is some evidence in favour of this view (Eysenck, 1961; see also Siegman, 1963; Mehryar, 1970), and the recent work done by the writer on sexual attitudes has shown that there are quite high correlations between permissive attitudes and sex practices, on the one hand, and extraverted personality features, on the other (Eysenck, 1970). Much more work needs to be done in relation to this hypothesis, of course, but if results should be confirmatory then we would seem here to have another reason for preferring the R and T solution to the Ferguson one.

On this view the 'authoritarian personality' would then simply correspond to a toughminded conservative, as suggested already in *The Psychology of Politics*; the 'humanitarian' would be the tenderminded radical. (The concept of the authoritarian personality has in recent years been criticized because of the alleged response set induced in respondents due to the uniformly positive way of phrasing the questions. Our factor I contains items phrased both positively and negatively, yet the factor emerges very much as it has done in earlier studies with the F scale; this suggests that the objections made on this account do not have much force.) These suggestions are made in full realization that preference for one solution or the other is based entirely on heuristic grounds; descriptively both solutions are essentially equal in value, and the alleged superiority of the one or the other must rest on the degree to which it suggests further research, and is supported by such research. It is clearly not possible at this writing to come to any definite conclusions on this point; only further research specially geared to testing the hypothesis proposed here (such as the personality–attitude correlation between T and E) can suggest an answer.

I am indebted to Mr David Newell and the B.B.C. for permission to use data collected by their Audience Research Department; needless to say, all the conclusions drawn are my own and do not commit the B.B.C. in any way.

REFERENCES

EYSENCK, H. J. (1947). Primary social attitudes. I. The organization and measurement of social attitudes. *Int. J. Opin. Attit. Res.* **1**, 49–84.

EYSENCK, H. J. (1951). Primary social attitudes as related to social class and political party. *Br. J. Sociol.* **2**, 198–209.

EYSENCK, H. J. (1954). *The Psychology of Politics*. London: Routledge & Kegan Paul.

EYSENCK, H. J. (1961). Personality and social attitudes. *J. soc. Psychol.* **53**, 243–250.

EYSENCK, H. J. (1970). *Psychology is about People*. London: Allen Lane Press.

FERGUSON, L. W. (1939). Primary social attitudes. *J. Psychol.* **8**, 217–223.

KAISER, H. F., HUNKA, S. & BRIANCHINI, J. (1969). Relating factors between studies based upon different individuals. In H. J. Eysenck & S. B. G. Eysenck (eds.), *Personality Structure and Measurement*. London: Routledge & Kegan Paul.

MEHRYAR, A. H. (1970). An attempt to cross-validate Eysenck's hypothesis regarding the relationship between extraversion and tough-mindedness. *J. soc. Psychol.* **80**, 109–110.

SIEGMAN, A. W. (1963). A cross-cultural investigation of the relationship between introversion–extraversion, social attitudes and antisocial behaviour. *Br. J. soc. clin. Psychol.* **2**, 196–208.

Manuscript received 28 *April* 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 213–220
With 3 text-figures
Printed in Great Britain

The Validity of the Eysenck Personality Inventory Studied by a Technique of Peer-rating Item by Item, and by Sociometric Comparisons

By H. B. GIBSON

Institute of Psychiatry, University of London

The validity of the scales of the Eysenck Personality Inventory was assessed by administering it to an undergraduate population (i) in the usual manner, and (ii) as though it were a rating instrument, with each student rating a classmate. The results showed significant correlations between self-rating on the N and E scales and ratings given by another. This was not so for the L scale; students tended to 'project' their own level of L on to the classmate they rated. Sociometric data were also obtained, and sociometric status was found to relate significantly to E but not to N. The personality characteristics of choosers in relation to those they chose were investigated.

A number of studies have been published by the Eysencks (Eysenck, 1962; Eysenck, S. B. G. & Eysenck, 1963, 1964) concerning the validity of the scales of extraversion (E) and neuroticism (N) as appearing in the Maudsley Personality Inventory (MPI) and the Eysenck Personality Inventory (EPI) and a larger inventory of items relating to the published scales. In assessing validity, the *method of nominated groups* was employed; this is a version of what Eysenck (1954) refers to as *agreement with ratings based on personal knowledge*. While this method of nominating groups especially high and especially low on the relevant scales establishes the relevance of the scales at the extremes it leaves some doubt as to the status of the more normal majority. In inviting raters to concentrate on abnormal personalities we are not validating the scales in a normal population; we are in fact dealing in 'ideal types' rather than in unidimensional scales. It has been shown that certain anomalies exist at the extreme of scales; e.g. those specially low on E on the MPI are high on N (Eysenck, 1959) and those specially high on L are low on N (Gibson, 1962).

Another technique of validating questionnaires by means of ratings based on personal knowledge is to use the same form of questionnaire both as a self-rating instrument (in the normal manner) and as an instrument for rating another person. In the latter case the rater is asked to judge the ratee on a large number of separate bits of behaviour—as many as there are items in the inventory. The rater may therefore be completely naive psychologically. He does not need to have any clear concept of the meaning of terms such as 'extraversion' and 'neuroticism' and indeed naivety may be a positive advantage as there will be less contamination by the rater being influenced by the face validity of the items (e.g. 'I think of Fred as being an extravert, therefore I expect he *really* would like to play pranks on others'). This approach to the problem of assessing validity is thus in direct contrast to the Eysencks' studies cited earlier and to those of Vingoe (1966) and Harrison & McLaughlin (1969) using self-rating, which have depended on the sophistication of judges being able to assess degrees of extraversion and neuroticism.

In the literature concerning the validity of measures of personality there is surprisingly little reference made to sociometric techniques. While it is generally assumed that extraversion is associated with social popularity there appears to be rather a dearth of experimental evidence supporting the idea. Northway (1967) reviews some of the studies associating personality variables with sociometric status, but these depend mainly on the results of projective tests with school children. These indicated that the children with most stable personality adjustment tended to be intermediate in sociometric status. Richards (1967) used the Bristol Social Adjustment Guide and found that children with the highest sociometric status were those with the highest personality stability. Holden (1969) used the Junior EPI and found that extraverts had higher sociometric status than introverts, stable extraverts being best able to make a public show. Holden's study is of considerable value in exploring the relationships between sociometric status and scores on the Junior EPI as it was conducted under various conditions of school organization. Neuroticism was shown to be associated with low sociometric status, the neurotic introvert being particularly liable to melt into the social background in competitive, streamed classes.

The following study was designed to investigate the validity of the three scales of the EPI in a student population by means of comparing the usual self-rating scores with scores obtained by using the inventory as though it were an instrument for rating other people. Secondarily, validity was assessed by relating the scores to objective measures of sociometric status.

METHOD

The EPI (Form A) was administered in the normal manner to 38 undergraduate students in a Humanities Department. These students habitually worked in four classes and so were well known to their own classmates. At the same time each student was asked to make a list ranking his classmates in order of how well he knew them. On a later occasion, each student was given an EPI form and asked to complete it as though it were a rating inventory concerning another member of his class nominated by the experimenter. The nominations were chosen so that each student was rating someone who appeared among the first three names he had listed, i.e. someone especially well known to him. It was stressed that the rating should be *objective*, i.e. the items should be answered in terms of how the ratee habitually appeared actually to behave or to feel, and not in terms of how he might be supposed to respond to the item. Thus if the rater thought that his friend would probably cheat the customs if he could, he should say so, irrespective of whether he thought that his friend would admit to doing it.

EPI. The forms were scored with the standard keys. For each subject three sets of scores were thus obtained. 'Subjective' E_s, N_s, L_s: the ordinary scores obtained by self-rating. 'Objective' E_o, N_o, L_o: the scores which the rater had awarded him. 'Projective' E_p, N_p, L_p: the scores which derived from his rating of a classmate. It should be noted that the labels are used for convenience only, and have no theoretical implications regarding subjectivity, objectivity or projection.

Sociometric scores. Each subject was awarded a sociometric score on the basis of his position in the rankings of his classmates. The first four scores only on each list were considered, and marks awarded: 1st place = 4, 2nd = 3, 3rd = 2, 4th = 1. Marks were summated for each individual, and then his percentile rank within his own class was calculated. The sociometric data obtained appeared to be of the usual sort, i.e. within each group there were certain 'stars' whom many people claimed to know well, and certain near 'social isolates' whom few people ranked high on their lists.

Table 1. *EPI data obtained under two conditions of administration compared with student norms*

	E		N		L	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Subjective	12·289	4·046	9·921	4·999	2·132	1·630
Objective and projective	12·947	5·784	8·789	5·443	2·132	1·934
Normative (Eysenck, H. J. & Eysenck, 1964)	11·095*	4·453	10·006*	5·006	2·263†	1·572

* Students. † Total population.

RESULTS

Table 1 shows the means and standard deviations of the three scales under the two conditions of administration, together with the normative data for students given by Eysenck, H. J. & Eysenck (1964). It will be observed that the means and standard deviations of the O and P conditions are of course identical because the 'objective' rating on one student is the 'projective' rating of another.

Table 2 gives the intercorrelations among all the personality scales and sociometric status. Again it will be noted that the three within-condition *r*'s of the 'objective' measures are identical with those of the 'projective' measures. This matrix was analysed further by the method of McQuitty (1957), which divided it into three separate groups: (i) $E_s \rightleftharpoons E_o \leftarrow \text{Status} \leftarrow E_p$; (ii) $L_s \rightleftharpoons L_p \leftarrow N_p$; (iii) $N_s \rightleftharpoons N_o \leftarrow L_o$.

(i) The strongest relationship in the matrix is seen to be between E_s and E_o ($r = 0·614$); in other words, the subjective scores of extraversion are very strongly validated by reports on individual items of behaviour made by a rater. Associated with both these scores is sociometric status; the more extraverted students are rated as being significantly better known in their groups. It is of interest to note that status is rather more associated with E_o (0·581) than with E_s (0·358). In order to investigate this relationship further, the status scores were ranked into six groups and the average scores on E_s and E_o were calculated for each group. The results

Table 2. *Intercorrelations among the personality scales and sociometric status*

(Decimal points omitted)

		1	2	3	4	5	6	7	8	9
1	N_o	—								
2	E_o	−089								
3	L_o	−485**	−227							
4	N_s	441**	−075	−245						
5	E_s	220	614***	−150	028					
6	L_s	−070	−005	097	−121	158				
7	N_p	095	192	−023	299	371*	−027			
8	E_p	164	−030	−227	061	154	−125	−089		
9	L_p	−028	−251	002	−024	−257	534***	−485**	−227	
10	Status	−166	581***	−003	−074	358*	−166	154	327	−337*

* $P < 0·05$. ** $P < 0·01$. *** $P < 0·001$.

are shown in Fig. 1. The interesting feature of this graph is that the subjects least well known in the group (those below the 10th percentile on status) are *not* the most introverted but are average on E.

It may be seen that the smallest r in this first McQuitty group is that of E_p with status.

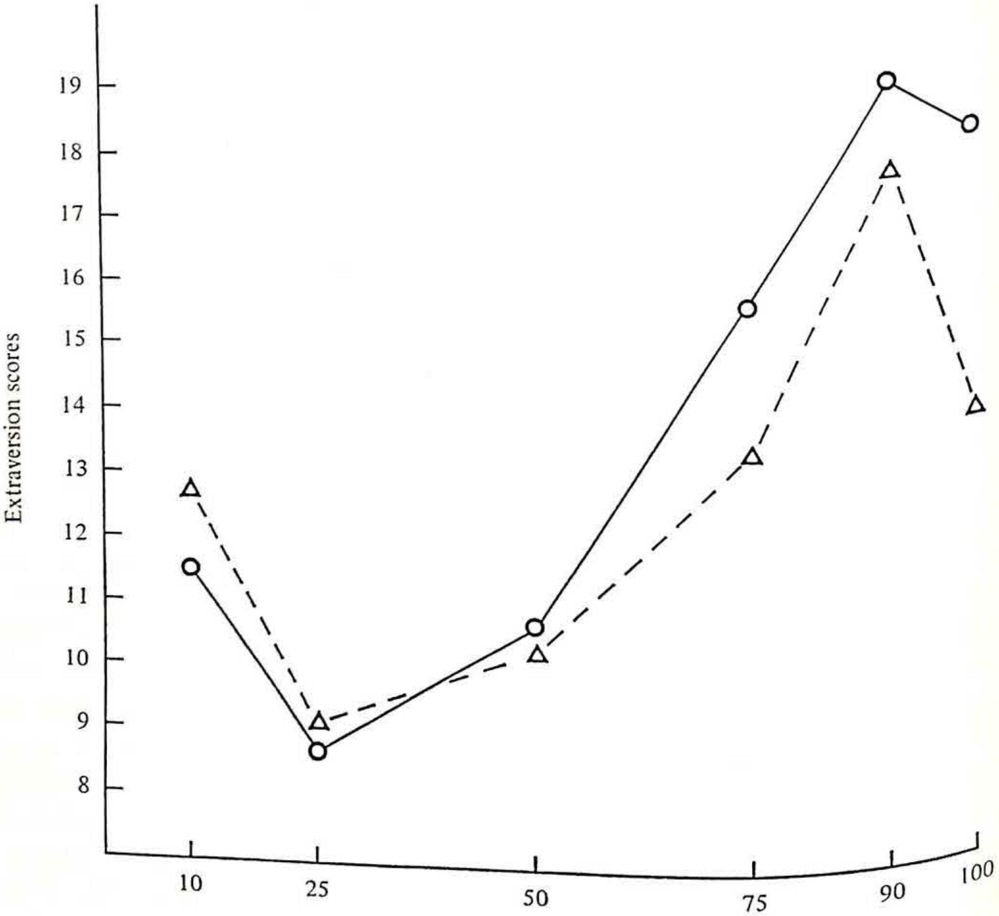


Fig. 1. Association between sociometric status and level of extraversion. o——o, E_s; △---△, E_o.

(ii) The next group concerns the relationship of L_s to L_p ($r = 0{\cdot}534$). Those who idealize themselves and obtain a high L score are also likely to idealize those whom they rate in terms of L scale items. Idealizing the person rated is also negatively associated with the N_p score which results from the rater's report on the N items ($r = -0{\cdot}485$).

(iii) This group concerns the relationship between N_s and N_o ($r = 0{\cdot}441$). Self-report on neuroticism is validated by the rater's report on the individual items which make up the scale. Again the feature noted above makes L_o negatively associated with N_o.

It is to be noted that the McQuitty analysis groups together all the variables which are associated by statistically significant r's, with one exception—that of status with L_p ($r = -0{\cdot}337$).

Personality characteristics determining sociometric choice

It is of interest to see the relationships between chooser and chosen in terms of personality characteristics. For this purpose the measures of N_s and E_s only were used. The mean of the scores of the first three choices made by each subject was taken as representing the personality type whom he chose.

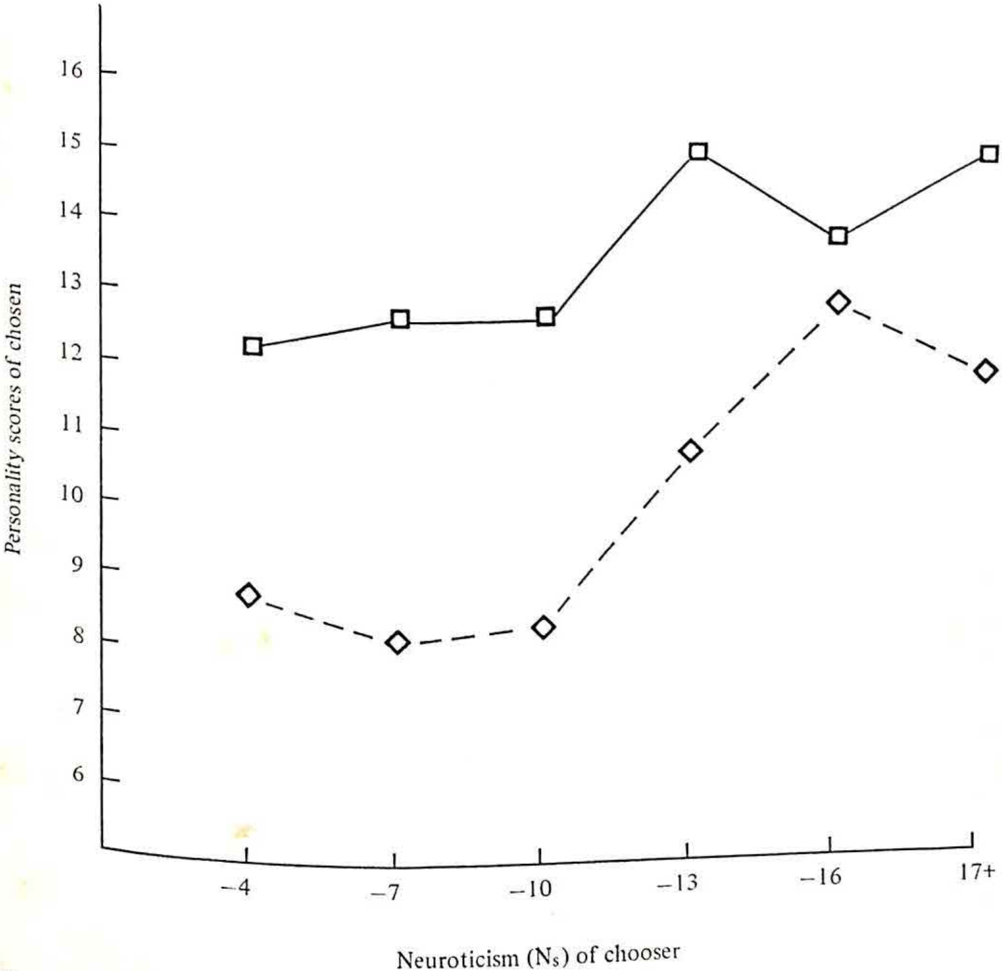


Fig. 2. Association between the level of neuroticism of the chooser and the personality scores of the chosen. □——□, E_s; ◇---◇, N_s.

All subjects were ranked in terms of their N_s scores and these were divided into six arrays as shown on the abscissa of Fig. 2. The mean personality scores of the classmates chosen were then plotted against the means of these arrays. As shown, those subjects with N scores > 10 chose friends who were significantly more extraverted ($\chi^2=6{\cdot}78$; d.f. $=1$; $P<0{\cdot}01$) and significantly more neurotic ($\chi^2=7{\cdot}73$; d.f. $=1$; $P<0{\cdot}01$).

A similar analysis was done concerning the level of E_s of the choosers, as shown in Fig. 3. Here there is a slight but non-significant positive relationship between the level of E of chooser and chosen, and a significant non-linear relationship between the E scores of the chooser and the N scores of the chosen. Introverts with E scores <10 (nine subjects) tend to choose friends with low N scores (by Fisher's test, $P=0{\cdot}003$).

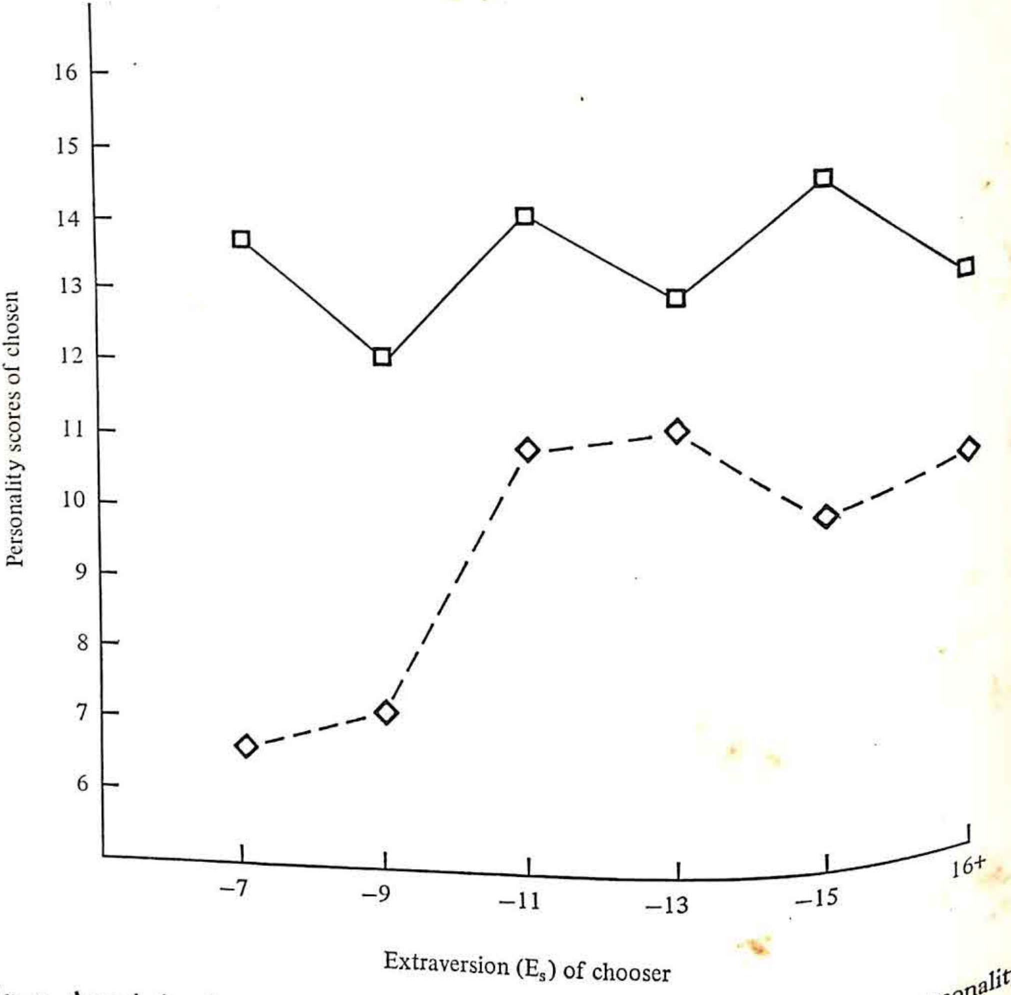


Fig. 3. Association between the level of extraversion of the chooser and the personality scores of the chosen. □——□, E_s; ◇---◇, N_s.

DISCUSSION

The major finding of this study is that rating a friend on the 48 items of behaviour which make up the N and E scales of the EPI produces scores which are significantly correlated with the scores that person obtains by completing the inventory in the usual manner. It was found that the correlation between E_o and E_s was rather higher than that between N_o and N_s. This may mean that it is easier to rate another person on the items of behaviour which make up the E scale than on those making up the N scale, since people naturally tend to conceal neurotic manifestations even from their close friends. The work of the Eysencks and others on the validity of the N and E scales by the use of nominated groups is thus supported by the results obtained with the present technique.

The results obtained with the L scale are of a different nature. The correlation between L_o and L_s was insignificant ($r=0{\cdot}097$). The tendency is for the rater to 'project' on to the person he is rating his own level of L, thus the r between L_s and L_p is quite large ($r=0{\cdot}534$). The L score, 'subjective' and 'projective', is also significantly and negatively related to how a person reports on his own neuroticism and that of the person he is rating. Otherwise the 'projective' scores are not strongly associated with the other personality scores. This is in line with a finding by Eysenck, S. B. G. & Eysenck (1964) that the personality of judges is not significantly associated with their judgement of other people's level of N and E.

The sociometric data yielded the expected finding that extraversion is positively associated with social popularity. It should be noted that in this study the sociometric data are not of the usual kind. Generally, subjects are asked to nominate people whom they like best, or regard as their best friends; however, these subjects were asked to rank their classmates in order of *how well they knew them*. It is possible, therefore, that extraversion correlates more with 'knowability' than 'likeability'. However, the detailed sociograms which were drawn of each group did in fact show all the usual features of sociograms. The fact that sociometric status was associated with E_o even more strongly than E_s probably indicates that the aspect of a person's extraversion which is most obvious is his social extraversion. Fig. 1 shows that the group of students lowest in sociometric status were *not* the most introverted. This relates to the special circumstances at the university—these were mainly an out-group of moderately extraverted students whose keen participation in extracurricular activities made them rather strangers to the working group.

Contrary to the findings cited concerning school children, the correlation of neuroticism with status was quite insignificant. Some of the sociometric stars scored highly on both N_o and N_s. This is probably a feature of a student population where N represents high drive rather than any personal maladjustment. Status was negatively correlated with L_p ($r=-0{\cdot}337$), a finding which is not easy to explain. Fig. 2 shows that subjects with a higher level of neuroticism tend to choose friends who are themselves higher on both neuroticism and extraversion. The latter feature may relate to the greater empathy of more extraverted people who are perhaps more congenial companions for those who, in the words of an EPI item, 'Often need understanding friends to cheer (them) up'. Fig. 3 shows no marked association between the levels of extraversion of the choosers and the chosen, but introverts with E scores of < 10 tend to choose friends with low N scores. It seems that introverts tend to exclude those high on neuroticism from close acquaintance, possibly because they are less willing to be concerned with other people's troubles.

REFERENCES

EYSENCK, H. J. (1954). *The Psychology of Politics*. London: Routledge & Kegan Paul.

EYSENCK, H. J. (1959). *Manual of the Maudsley Personality Inventory*. London: University of London Press.

EYSENCK, H. J. & EYSENCK, S. B. G. (1964). *Manual of the Eysenck Personality Inventory*. London: University of London Press.

EYSENCK, S. B. G. (1962). The validity of a personality questionnaire as determined by the method of nominated groups. *Life Sci.* **1**, 13–18.

EYSENCK, S. B. G. & EYSENCK, H. J. (1963). The validity of questionnaire and rating assessments of extraversion and neuroticism, and their factorial stability. *Br. J. Psychol.* **54**, 51–62.

EYSENCK, S. B. G. & EYSENCK, H. J. (1964). The personality of judges as a factor in the validity of their judgments of extraversion-introversion. *Br. J. soc. clin. Psychol.* **3**, 141–144.

GIBSON, H. B. (1962). The lie scale of the Maudsley Personality Inventory. *Acta psychol.* **20**, 18–23.

HARRISON, N. W. & MCLAUGHLIN, R. J. (1969). Self-rating validation of the Eysenck Personality Inventory. *Br. J. soc. clin. Psychol.* **8**, 55–58.

HOLDEN, L. A. (1969). The effects of personality type and of changes in junior school streaming policies on goal-setting, social relations, and symbolic learning. (Unpublished Ph.D. thesis, University of London.)

MCQUITTY, L. L. (1957). Elementary linkage analysis. *Educ. psychol. Measur.* **17**, 207–229.

NORTHWAY, M. L. (1967). *A Primer of Sociometry*. Toronto: University of Toronto Press.

RICHARDS, H. H. (1967). Psychological factors associated with the sociometric status of children attending a comprehensive school in Breconshire. *Br. J. educ. Psychol.* **37**, 261–262.

VINGOE, F. J. (1966). Validity of the Eysenck extraversion scale as determined by self-ratings in normals. *Br. J. soc. clin. Psychol.* **5**, 89–91.

Manuscript received 28 *April* 1970

Revised manuscript received 15 *August* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 221–227
Printed in Great Britain

The Effects of Ordinal Position on Measured Intelligence and Peer Acceptance in Adolescence

By C. MURRAY

St Mary's College, Twickenham, Middlesex

Data are presented from a cross-sectional survey of third-year male pupils in 12 secondary modern schools and analysed with regard to the effect of ordinal position on measured intelligence and peer acceptance. In contrast to previous research findings no significant ordinal effects on the dependent variables are noted with the data classified as to age, sex, social status and family size.

Reviews of studies which have incorporated birth order as a research variable either in a restricted area (e.g. Schachter, 1963) or in a more general context (e.g. Altus, 1966; Warren, 1966) are readily available. In a further review the lack of a systematic theoretical consideration of the relationship between birth order and its various correlates has recently been discussed by Kammeyer (1967) and running throughout this and the other summaries is the notion that 'More careful study of birth order than has been demonstrated in the (48) studies reviewed...seems desirable. Correlates of birth order at all age levels...should be studied for an adequate understanding' (Warren, 1966, p. 48). The most recent endorsement of this view with regard to the necessity for greater control over possible confounding effects in birth order research is contained in Adams (1969).

From the reviews mentioned above it would appear that one of the most consistent correlates of ordinal position is that of intellectual aptitude as manifest in performance test score, eminence and/or college attendance. A summary derived from the review by Altus (1966) of some relevant findings with respect to this dependent variable is shown in Table 1.

Altus (1966, p. 48) concludes: 'It seems reasonable to believe that the aptitude data and the college attendance figures must be interrelated and it seems equally

Table 1. *Birth-order studies reviewed by Altus with 'intellectual aptitude' as dependent variable*

Author	Sample	Results
Terman (1925)	1000 gifted children (140+ IQ)	First born over-represented
Stewart (1962)	7000 grammar and secondary modern pupils	First born over-rep. in grammar
Capra & Dittes (1962)	Undergraduates	First born over-represented
Schachter (1963)	628 high-school students	First born higher GPA
Schachter (1963)	199 undergraduates	First born over-represented
Nichols	1168 gifted children (150+ IQ)	First born over-represented
Altus (1964)	1800 'select' undergraduates	First born higher verbal IQ
Altus (1965)	2500 'select' undergraduates	First born higher verbal IQ

reasonable that both sets of data are linked, quite possibly in a causal way, to the numerous data on eminence that have been presented.' In terms of socialization theory these results may be 'explained' in terms of parental attitudes and experience, differential amounts of parent–child interaction, availability of child models and effects of parental age (Clausen & Williams, 1963). Physiological sources associated with birth order effects are mentioned by Warren (1966) as being connected with the intrauterine environment of the mother and various perinatal influences. The atypical nature of the samples employed in the studies tabled above and the difficulties imposed by secondary analysis with respect to control of possible covariates suggest that further investigation of this dependent is warranted.

Schachter (1959) has been one of the more successful social scientists in terms of relating theoretical explanation to birth order influence on non-cognitive variables; more specifically, he has recently suggested (Schachter, 1964), drawing 'semi-intuitively' from experimental research on dependency, that 'when exposed to a standard fear-inducing situation, first borns responded more timidly and leaned on other people considerably more than did later borns—hardly an appealing long-run picture and one that would lead to the suspicion that first borns would be less popular than later borns' (1964, p. 455). Support for this contention is ingeniously derived from the work of Finneran (1958) and Haeberle (1958). Schachter's results show that over his 15 fraternity and sorority groups that comprise the sample drawn to test this proposition, first borns are in fact more popular (summed choice score on three-choice, two-criteria sociometric test) than later borns ($P < 0{\cdot}01$). The select nature of the sample, the inclusion of 'first born and only' as a single category and the lack of control in respect of family size limit the generalization of these important findings. It should also be mentioned that Murdock & Smith (1969) have recently pointed to the inconsistent series of results furnished by studies relating birth order and affiliation, and Jamieson (1969), in reporting an experiment on risk-taking, highlights the importance of family size in psychological research dealing with ecological variables.

The aims of this study are to ascertain whether the relationships between birth order and different classes of dependent variables, cognitive and non-cognitive, discussed above would be replicated within a sample of English adolescents when the effects of social status, family size and age were controlled. The necessity for exercising some control, however crude, over social status and family size when IQ is the dependent variable is readily apparent from the most cursory glance at the research literature (e.g. Scottish Mental Survey, 1953); similarly, social status has been found to be positively related to acceptability by several investigators in America (e.g. Bonney, 1944; Elkins, 1958; Epstein & Komorita, 1965), though a similar relationship was not found in a recent small-scale study ($n = 267$) in this country (Murray, 1969). Age effects on sociometric status are not clear (Blyth, 1960), though this is an important covariate of IQ. The limitations of the study lie in the fact that birth order effects, if one accepts the differential socialization view, are unlikely to operate independently of other familial aspects, such as sex of siblings and difference in siblings' age; no control is exercised over these structural aspects in this study.

METHOD

Subjects. The data for this paper are extracted from a cross-sectional study involving motor, intellectual and psychosocial characteristics of third-year adolescents in 42 forms located in 12 secondary modern school systems. The schools were chosen for geographical spread and, arising from the requirements of the wider study, varied in size and composition. Three schools were located in Liverpool, four in Manchester, two in Coventry and three in London; overall, five were mixed and seven single-sexed.

Incomplete data on 919 subjects were collected by the author and a party of testers with complete data on social status, family size, IQ and sociometric acceptance being secured for 753 subjects. Forty-four 'only' children were excluded from this number leaving a total number of 709 for analysis. The percentages of first and later borns making up the 709 cases are tabulated in Table 2: numbers are given in parentheses.

Table 2. *Percentages and numbers (in parentheses) of first and later born adolescents at four levels of family size*

	Family size				
Birth order	2	3	4	5	Total
First borns	38 (62)	40 (58)	29 (38)	18 (47)	30 (205)
Later borns	62 (100)	60 (88)	71 (94)	82 (222)	70 (504)
Total	100 (162)	100 (146)	100 (132)	100 (269)	100 (709)

Measures

Indices. Social status was derived from occupation of father and coded according to the Hall–Jones classification (Hall & Jones, 1950) into eight categories; these categories were then further reduced to two, upper and lower 'classes'. Within each 'class' group a further classification was imposed in respect of four divisions of family size, viz. 2, 3, 4, 5+, and within each family size 'first borns' were separated from 'later borns'.

Dependent variables. IQ was measured by the AH 4 test devised by Heim (1967), which gives two measures, verbal and non-verbal IQ. The correlation between both these measures for the sample in this study ($n=709$) is 0·708. An index of peer acceptance was derived for each individual from the summed scores on a three-criteria (associate for gym, playground, library) three-choice sociometric test; the summed score being divided by the form $N-1$ and converted through percentiles to a normalized standard score (T score) with a mean of 50 and σ of 10. A second measure of peer acceptance, termed peer prestige, was calculated from a seven-point social distance scale described by Bogardus (1933), in which each form member rated every other. The score for each subject was his total weighted score (sum of squares of the scale values) divided by the number in the form with division being carried to one decimal place and the point dropped: low scores indicate high acceptance. The two sociometric instruments tap somewhat different facets of peer relations as indicated by the moderate correlation of $-0{\cdot}498$ ($n=709$) between the two measures for the sample used in this study. The use of a social distance scale, whereby the whole group is ranked in terms of social acceptance to give a measure of 'peer prestige', is quite laborious in terms of computational effort but would seem to be worthwhile in that it produces a more sensitive index of perceived individual status when combined with the more simply scored sociometric test (Bjerstedt, 1956). The two instruments differ chiefly in that the latter is concerned with choice of associate for a specific situation, with all that this implies, whereas the personal distance scale taps a generalized measure of 'psychological' acceptability; consequently we would not necessarily expect a very high correlation between scores given by these instruments (see Semler, 1960). The sociometric indices were so constructed that individual scores arising from forms of varying size are comparable and all the dependent variables as described are normally distributed.

This survey thus represents a non-orthogonal design with the dependent variables cross-classified by the factors of social status, family size and birth order. Tsao (1946), Yates (1934) and Stevens (1948) have suggested methods of meeting the problem of non-orthogonality, i.e. disproportionate frequencies in each subcell of a cross-classification and a program has been written which involves the method of fitting constants to sets of data (Simpson, 1964). However, a very flexible program has been recently developed (Baughan, 1970), whereby a linear model is utilized to represent the various Anova factors with allowed levels on each and which handles both the univariate and multivariate analysis of variance by obtaining a least-squares solution from which terms are successively removed by way of multiple-regression analysis. In that a conservative estimate of significance is given by this procedure, this method was preferred and the data prepared accordingly.

RESULTS

As the sample for this study is really a subsample of that for which incomplete data were available, we present the means and standard deviations of the dependent variables for both samples to check on any appreciable differences that may have occurred as a result of our selection procedure. The results (Table 3) indicate that our data closely represent those from which they are drawn.

Table 3. *Means and standard deviations of dependent variables using all possible observations and sample observations*

Dependent variables	No. of observations		Mean		S.D.	
	Possible	Sample	All	Sample	All	Sample
Verbal IQ	820	709	25·368	25·440	9·596	9·325
Non-verbal IQ	820	709	34·021	34·131	10·823	10·728
Peer acceptance	890	709	49·998	50·867	10·176	9·842
Peer prestige	908	709	19·981	19·730	6·552	6·498

The total *n* of 753 from which the 709 first and later borns were selected reflects the number of cases for which data on all four dependent variables and classification indices are available. Thus though we have 820 cases for IQ and 890 for peer acceptance only 753 of these provide data on IQ, peer prestige, social status and family size. Motor performance data were secured for all 919 subjects.

The major results of the study are brought together in Table 4.

As an aid to the interpretation of these results the data description was reformulated and the analysis re-run without control over social status. The results of this two-way reanalysis are shown in Table 5.

DISCUSSION

The results (Table 4) indicate that for each level of family size, with some control exercised over social status, there is no significant difference between first and later borns on the two intelligence or peer acceptance measures. Indeed, for the three dependent variables of peer acceptance, peer prestige and non-verbal IQ the *F* ratios approach zero. Interestingly, the *F* ratio of 3·68 with verbal IQ as dependent variable approaches significance at the 0·05 level, and inspection of the means (first borns 28·17, later borns 26·27) indicates the difference lies in the direction suggested by our previous review. In that verbal, as compared to non-verbal, IQ has been shown to be more 'affected' by environmental factors there is some slight support here for explanation of ordinal effects in terms of social rather than physiological influences.

Table 4. *Analysis of variance: the effects of social class, family size and ordinal position on cognitive and non-cognitive dependent variables*

Source of variation	S.S.	D.F.	M.S.	*F*	Significance level
Verbal IQ					
Classes	610·3072	1	610·3072	7·80	0·01
Sizes	2230·2920	3	743·4307	9·51	0·01
Positions	288·0529	1	288·0529	3·68	n.s.
C × S	776·7056	3	258·9019	3·31	0·05
C × P	1·9088	1	1·9088	0·02	n.s.
S × P	474·5437	3	158·1812	2·02	n.s.
C × S × P	201·4738	3	67·1579	0·86	n.s.
Residual	54111·6205	693	78·1960		
Non-verbal IQ					
Classes	521·5250	1	521·5250	4·95	0·05
Sizes	2985·8858	3	995·2953	9·45	0·01
Positions	18·1844	1	18·1844	0·17	n.s.
C × S	514·9917	3	171·6639	1·63	n.s.
C × P	8·3433	1	8·3433	0·08	n.s.
S × P	375·6358	3	125·2119	1·19	n.s.
C × S × P	73·5431	3	24·5114	0·23	n.s.
Residual	72850·1313	693	105·2748		
Peer acceptance					
Classes	109·6313	1	109·6313	1·13	n.s.
Sizes	201·4111	3	67·1370	0·69	n.s.
Positions	42·9546	1	42·9546	0·44	n.s.
C × S	143·8286	3	47·9429	0·49	n.s.
C × P	7·3396	1	7·3396	0·07	n.s.
S × P	511·3041	3	170·4347	1·76	n.s.
C × S × P	346·4543	3	115·4848	1·19	n.s.
Residual	67089·6795	693	96·9504		
Peer prestige					
Classes	113·8595	1	113·8595	2·70	n.s.
Sizes	170·6990	3	56·8997	1·35	n.s.
Positions	0·0079	1	0·0079	0·00	n.s.
C × S	63·5102	3	21·1701	0·50	n.s.
C × P	14·1371	1	14·1371	0·34	n.s.
S × P	154·5079	3	51·5026	1·22	n.s.
C × S × P	84·8859	3	28·2953	0·67	n.s.
Residual	29145·0734	693	42·1172		

When no control is exercised over social status (Table 5) we note a significant difference between first and later borns ($P < 0·05$) in performance on verbal IQ, first borns obtaining a higher mean value. In the light of the results discussed above this finding would seem to be largely artifactual, arising from the lack of control over social status effects; again no significant differences are noted with regard to the other three dependent variables.

Though the sample, in terms of intellectual attributes, is 'representative' of approximately 80 per cent of the school-going population of this age, it is to an extent 'downwardly selected' for IQ, whereas the previous work already described has taken place largely within groups which exhibit some degree of 'upward selection' as far as intelligence is concerned. We may speculate that this sampling feature and the partial control of social status effects may explain the negative findings of this investigation when other studies have obtained positive results. With regard to the cognitive dependent variables, it was not the intention of this study to consider the extent of 'representation' of first borns within our sample population; however, arising from our considerations of the relationship between ordinal position and IQ we might expect that first borns would be under-represented in our sample, in that the brightest children have been selected out at 11+. We have no normative data for percentages of first and later borns in the general population of adolescents but Table 2 of this report indicates that marked under-representation has not occurred, with just over a third of the total sample being classified as first born.

Table 5. *Analysis of variance: the effects of family size and ordinal position on cognitive and non-cognitive dependent variables*

Source of variation	S.S.	D.F.	M.S.	*F*	Significance level
Verbal IQ					
Sizes	4330·8079	3	1443·6026	18·16	0·01
Positions	363·7594	1	363·7594	4·58	0·05
S × P	559·7459	3	186·5820	2·35	n.s.
Residual	55718·5262	700	79·5971		
Non-verbal IQ					
Sizes	6216·4495	3	2072·1498	19·56	0·01
Positions	14·9765	1	14·9765	0·14	n.s.
S × P	452·8024	3	150·9341	1·42	n.s.
Residual	74275·9568	700	106·1071		
Peer acceptance					
Sizes	306·8281	3	102·2760	2·43	n.s.
Positions	5·4891	1	5·4891	0·13	n.s.
S × P	223·2969	3	74·4323	1·77	n.s.
Residual	29492·3576	700	42·1314		
Peer prestige					
Sizes	106·3031	3	35·4344	0·35	n.s.
Positions	62·5255	1	62·5255	0·63	n.s.
S × P	547·2393	3	182·4121	1·83	n.s.
Residual	69913·7985	700	99·0185		

As an incidental observation we note that, perhaps as a consequence of some under-representation of the 'upper' social status groups, family size appears to have a more pronounced effect on IQ than social status, particularly in the case of non-verbal IQ. As expected, high social status and small family size are significantly associated with IQ, though not with measures of sociometric acceptance. This latter result runs contrary to findings reported in the American literature and supports the incidental results of the previous small-scale study already referred to (Murray, 1969).

REFERENCES

ADAMS, B. N. (1969). Birth order and college attendance: a re-evaluation of a re-evaluation. *Sociometry* **32**, 503–504.

ALTUS, W. D. (1966). Birth order and its sequelae. *Science, N.Y.* **151**, 44–49.

BAUGHAN, B. (1970). Program G.U.M.A.—multivariate analysis of variance with canonical variates. (Unpublished memorandum, University of London.)

BJERSTEDT, A. (1956). *Interpretations of Sociometric Choice Status*. Lund: Gleerup.

BLYTH, W. A. L. (1960). The sociometric study of children's groups in English schools. *Br. J. educ. Stud.* **3**, 127–147.

BOGARDUS, E. S. (1933). A social distance scale. *Sociology sociol. Res.* **17**, 265–271.

BONNEY, M. E. (1944). Relationship between social success, family size, socio-economic home background and intelligence. *Sociometry* **7**, 26–37.

CLAUSEN, J. A. & WILLIAMS, J. R. (1963). Sociological correlates of child behaviour. In H. W. Stevenson (ed.), *Child Psychology*. Chicago: University of Chicago Press.

ELKINS, D. (1958). Some factors related to the choice status of ninety, eighth-grade children in a school society. *Genet. Psychol. Monogr.* **58**, 207–272.

EPSTEIN, R. & KOMORITA, S. S. (1965). Parental discipline, stimulus characteristics of out-groups and social distance in children. *J. Person. soc. Psychol.* **2**, 416–420.

FINNERAN, M. P. (1958). Dependency and self concept as functions of acceptance and rejection by others. *Am. Psychol.* **13**, 332. (Abstract.)

HAEBERLE, A. (1958). Interactions of sex, birth order and dependency with behaviour problems and symptoms in emotionally disturbed pre-school children. (Paper read at Eastern Psychological Association.)

HALL, J. & JONES, D. C. (1950). The social grading of occupations. *Br. J. Sociol.* **1**, 31–55.

HEIM, A. W. (1967). *A.H.4 Group Test of General Intelligence*, rev. ed. Slough: N.F.E.R.

JAMIESON, B. D. (1969). The influences of birth order, family size and sex differences on risk-taking behaviour. *Br. J. soc. clin. Psychol.* **8**, 1–8.

KAMMEYER, K. (1967). Birth order as a research variable. *Social Forces* **46**, 71–80.

MURDOCK, P. H. J. & SMITH, G. F. (1969). Birth order and affiliation. *Br. J. soc. clin. Psychol.* **8**, 235–245.

MURRAY, C. (1969). A study of physical and sociometric measures of adolescents occupying differential academic and social status positions. (Unpublished M.Phil. thesis, University of London.)

SCHACHTER, S. (1959). *The Psychology of Affiliation*. Stanford: Stanford University Press.

SCHACHTER, S. (1963). Birth order, eminence and higher education. *Am. sociol. Rev.* **28**, 757–768.

SCHACHTER, S. (1964). Birth order and sociometric choice. *J. abnorm. soc. Psychol.* **68**, 453–456.

SCOTTISH MENTAL SURVEY (1953). *Social Implications of the Scottish Mental Survey*. London: University of London Press.

SEMLER, J. J. (1960). Relationships among several measures of pupil adjustment. *J. educ. Psychol.* **51**, 60–64.

SIMPSON, H. R. (1964). The influence of home and school on test performance. In J. W. B. Douglas, *The Home and the School*. London: MacGibbon & Kee.

STEVENS, W. L. (1948). Statistical analysis of a non-orthogonal tri-factorial experiment. *Biometrika* **35**, 346–367.

TSAO, T. (1946). Analysis of variance with disproportionate cell frequencies. *Psychometrika* **11**, 122–124.

WARREN, J. R. (1966). Birth order and social behaviour. *Psychol. Bull.* **65**, 38–49.

YATES, F. (1934). The analysis of multiple classifications with unequal numbers in the different classes. *J. Am. statist. Ass.* **29**, 51–56.

Manuscript received 29 *May* 1970

Revised manuscript received 20 *November* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 228–233
Printed in Great Britain

Expectancy Confirmation and Affectivity

By ROGER L. TERRY

Department of Psychology, Hanover College

Situations involving the confirmation and disconfirmation of expectancies varying dichotomously in strength and value are theorized to involve three components: outcome, strength and value, and to be ordered in terms of positive affect generated. Results of an experiment closely parallel the theoretically derived order, indicating that confirmation generates greater positive affect than disconfirmation; the positive affect generated by confirmation (disconfirmation) is directly (inversely) related to expectancy strength; and confirmation of positive expectancies generates greater positive affect than confirmation of negative expectancies; and the obverse of this.

Expectancy theory is concerned with the affective reactions of persons to the confirmation and disconfirmation of their expectancies. Situations involving expectancy confirmation contain at least two sources of affect—affect due to confirmation and disconfirmation, *per se*, and affect due to the obtained outcome. For convenience, the former source is said to generate cognitive affect, and the latter is said to generate hedonic affect.

Predictions of hedonic affect are generally straightforward. The occurrence of positively valued outcomes and the non-occurrence of negatively valued outcomes tend to generate greater positive affect than the non-occurrence of positively valued outcomes and the occurrence of negatively valued outcomes.

One would readily predict that if the expected event is positively valued, confirmation is preferred; if the expected event is negatively valued, disconfirmation is preferred. However, research by Day (1961) on positively valued outcomes and by Festinger *et al.* (1956) on negatively valued outcomes indicates that sources of cognitive and hedonic affect interact. Specifically, the effects of value of the expected event are theorized to be inversely related to the effects of expectancy strength.

With respect to cognitive affect, we reasoned that if the expectancy held is relatively strong, one is more motivated to maintain it than to acquire a new one to replace it. Thus confirmation of strong expectancies should generate greater positive affect than disconfirmation. Similarly, if the expectancy held is relatively weak, one is more motivated to replace it than to maintain it. Thus disconfirmation of weak expectancies should generate greater positive affect than confirmation. In sum, positive cognitive affect generated by confirmation (disconfirmation) is directly (inversely) related to the strength of the expectancy involved.

Now, a person holding a strong expectancy 'expects' confirmation, while a person holding a weak expectancy 'expects' disconfirmation. Therefore confirmation of a strong expectancy involves the occurrence of an event consistent with the person's expectancy, while disconfirmation involves the occurrence of an event discrepant from his expectancy. Likewise, confirmation of a weak expectancy involves the occurrence of a discrepant event, while disconfirmation involves the occurrence of a consistent event.

Thus affectivity is viewed as a function of three dichotomized components: (1) *outcome*, with elements consistent and discrepant; (2) *strength*, with elements strong and weak; and (3) *value*, with elements positive and negative. If it is assumed, following Bramel (1968, p. 358), that the effects of value are not extreme, it can be shown that the Cartesian product of the elements of these three components is ordered in terms of degree of positive affect generated, as shown in Table 1.

Table 1. *Facet design and descending order of expectancy variables in terms of positive affect generated*

	Components			
	Outcome	Strength	Value	Operations
I	Consistent	Strong	Positive	Occurrence of an expected positive event
II	Consistent	Strong	Negative	Occurrence of an expected negative event
III	Consistent	Weak	Negative	Non-occurrence of an unexpected negative event
IV	Consistent	Weak	Positive	Non-occurrence of an unexpected positive event
V	Discrepant	Weak	Positive	Occurrence of an unexpected positive event
VI	Discrepant	Weak	Negative	Occurrence of an unexpected negative event
VII	Discrepant	Strong	Negative	Non-occurrence of an expected negative event
VIII	Discrepant	Strong	Positive	Non-occurrence of an expected positive event

A 2×2×2 factorial experiment, involving the dichotomous and orthogonal manipulation of outcome, strength and value was designed to test these notions. The eight experimental conditions were predicted to be ordered in terms of positive affect generated according to Table 1. A significant outcome main effect, outcome × strength interaction, and second-order interaction were all therefore predicted.

PROCEDURE

Eighty, paid, introductory psychology students volunteered to participate in one of eight sessions of a study ostensibly designed to investigate the effects of hunger on food preferences. When signing up for a session, the subject read that he was not to eat for 5 hours prior to attending; all sessions were conducted at 5 p.m. The subjects arrived in groups of 10 and were led through a large room, set up to give the impression that food was to be served, into the actual experimental room. Thus all subjects arrived strongly expecting to taste some foods.

Immediately, weak expectancies of eating in the weak expectancy (E:W) conditions were created. Before the experimenter introduced the project to these subjects, they were informed that the experimenter's assistant could not obtain the food supplies and that while they would not taste any foods, they would fill out some forms to be used for control purposes.

In all conditions, the experimenter stated an interest in the effects of hunger and order of presentation on food preferences. This discussion was to add credibility to the ostensible purpose of the study, and to allow the experimenter to emphasize the importance of control procedures in experiments, thus justifying the request that subjects in the E:W groups respond to several forms for purposes of experimental control.

Value of the expected event was then manipulated by informing subjects to be in the positively valued expectancy (V:P) groups that the foods were quite enjoyable, while subjects to be in the V:N groups were informed that the foods were very unpleasant.

Following a brief interval, the experimenter explained that there were factors other than hunger and order effects which influence food preferences and that he would distribute two sets of questions designed to assess these other factors.

The first set of questions asked for such information as the subject's age, sex, height, weight, favourite meat, etc. The subject also indicated how pleasant the task was to be.

At the completion of these questions, outcome was orthogonally manipulated. In half of the E:S groups, the experimenter distributed a two-page set of questions saying that after it was completed, the subject was to go into the outside room to obtain his eating utensils and food supplies. Thus the subject's strong expectancy of eating was confirmed with this consistent outcome. In the other E:S groups, the experimenter reported, after a staged telephone conversation, that a problem had come up and that there would not be enough time to sample the foods. Thus the subject's strong expectancy was disconfirmed with this discrepant outcome. (The subject was then requested to respond to the questions with the justification that the results would be used for control purposes.) In half of the E:W groups, the experimenter distributed the questions, repeating that while other subjects tasted the foods after answering these questions, the subject would not do it himself and that the results would be used for control purposes. This outcome was consistent with (i.e. disconfirmed) the subject's weak expectancy of eating. In the other E:W groups, the experimenter's assistant suddenly arrived with food supplies and the experimenter distributed the questions, saying that the subject would taste the food after all. This outcome was discrepant from (i.e. confirmed) the subject's weak expectancy.

When the subject had completed the two-page set of questions, the experiment was over, and the experimenter explained the real purpose of the study.

The two-page set of questions was designed to measure the three independent variables and several aspects of the dependent variable of affectivity. The following task was posed: 'On each of the following scales indicate how you felt when the experimenter informed you that you were all set to indicate your food preferences.' (For those groups in which food was not to be served, 'not' was substituted for 'all set'.) The scales were all six-point, Likert-type scales, containing no zero or neutral point.

For example, to assess the success to which the outcome and strength variables were manipulated, the subject's responses to the following alternatives were analysed: extremely dissonant, quite dissonant, dissonant, harmonious, quite harmonious, extremely harmonious. Another scale was constructed from the polar adjectives, surprised and not surprised.

Ten additional scales were included to index the dependent variable of affectivity (see Table 2). The direction of the scales was randomized to inhibit the formation of a response set; the subjects' responses were scored so that a high value would indicate high positive affect.

RESULTS

On the first set of questions distributed, the subject indicated on a 12-point scale how enjoyable he foresaw his task to be and how much he was looking forward to participating in the task. An analysis of the variance in the responses to both of these questions summed together revealed only the value main effect ($F=7{\cdot}762$; d.f. $=1$, 72; $P<0{\cdot}01$) and the value × strength interaction ($F=6{\cdot}793$; d.f. $=1$, 72; $P<0{\cdot}025$). The directions of these differences were in line with predictions if the value and strength variables were manipulated as intended.

The subject's responses to the dissonance and surprise scales were summed and an analysis of variance was performed on these data. The outcome main effect ($F=10{\cdot}937$; d.f. $=1$, 72; $P<0{\cdot}01$) and the outcome × strength interaction ($F=21{\cdot}028$; d.f. $=1$, 72; $P<0{\cdot}001$) were the only significant sources of variance. These results were also perfectly in line with predictions if the outcome and strength variables were manipulated as intended.

The mean results from the ten affectivity scales are presented in Table 2. Note that the higher a mean score, the higher the indexing trait of positive affect.

Analyses of the variance in each of these sets of means were performed; the F

Table 2. *Mean affectivity scores of the eight experimental groups*

	Experimental group							
	I	II	III	IV	V	VI	VII	VIII
Calm–Agitated	4·5	4·2	4·3	3·7	4·4	3·4	3·5	3·5
Light–Heavy	3·8	3·6	3·8	3·9	3·5	3·7	3·9	4·0
Pleased–Displeased	4·3	4·0	3·9	3·3	4·2	3·5	2·8	2·8
Happy–Disappointed	4·2	4·2	3·6	3·3	4·1	3·3	2·9	2·6
Friendly–Hostile	4·3	4·2	4·2	3·9	4·1	3·7	3·7	3·7
Grateful–Ungrateful	4·0	4·0	3·9	3·6	4·0	3·4	3·2	2·8
Warm–Cold	4·1	4·2	4·0	4·0	4·4	3·7	3·4	3·7
Relaxed–Tense	4·2	4·4	4·2	3·8	3·8	3·4	3·6	3·5
Positive–Negative	4·3	3·9	3·7	3·8	3·9	3·4	3·4	3·1
High–Low	4·0	3·9	3·7	3·7	3·6	3·6	3·0	3·5
Total positive affect	41·7	40·6	39·3	37·0	40·0	35·1	33·4	33·2

ratios for the several sources of variation in the affectivity scales, individually and summed, are presented in Table 3.

Examination of the analyses of each of the individual scales indicated that the prediction of an outcome main effect was supported three times at the 0·025 level of confidence, twice at the 0·01 level, and four times at the 0·001 level. The second prediction of an outcome × strength interaction was supported twice at the 0·05 level of confidence, twice at the 0·01 level, and three times at the 0·001 level. Finally, the prediction of a second-order interaction was supported once at the 0·05 level and twice each at the 0·025 and 0·01 levels of confidence. All of these differences were in the predicted direction.

(The only unpredicted findings from these individual analyses appeared in the responses to the warm–cold scale. The value main effect and the outcome × value interaction both obtained conventional levels of significance.)

Table 3. *F ratios (d.f. = 1, 72) for the mean affectivity scores*

Source	Calmness	Lightness	Pleasure	Happiness	Friendliness
Outcome (O)	6·776**	< 1	14·336****	17·226****	6·430**
Strength (S)	< 1	< 1	2·962	< 1	< 1
Value (V)	< 1	< 1	< 1	< 1	< 1
O × S	4·207*	1·956	30·332****	34·652****	2·100
O × V	3·114	< 1	2·962	1·918	1·181
S × V	< 1	< 1	< 1	1·918	< 1
O × S × V	6·776**	< 1	7·583***	5·875**	2·100

Source	Gratefulness	Warmth	Relaxation	Positivity	Height	Total positive affect
Outcome (O)	19·034****	6·255**	11·885***	12·878****	8·122***	27·775****
Strength (S)	3·483	2·510	< 1	< 1	< 1	< 1
Value (V)	< 1	4·156*	< 1	1·738	1·142	1·191
O × S	15·552****	8·683***	1·095	8·006***	4·568*	20·446****
O × V	1·034	6·172**	1·795	< 1	< 1	3·384
S × V	2·138	1·317	< 1	< 1	1·142	< 1
O × S × V	7·310***	< 1	1·113	4·331*	< 1	7·027***

* $P < 0{\cdot}05$. ** $P < 0{\cdot}025$. *** $P < 0{\cdot}01$. **** $P < 0{\cdot}001$.

To determine the closeness of fit between the theoretically derived order and the obtained order of the experimental groups in terms of *total* positive affect generated, Kendall's tau was computed, and found to equal 0·857, which was significant beyond the 0·001 level of confidence. The only deviation from the theoretical order was an inordinately high degree of positive affect generated in the condition involving the occurrence of an unexpected positive event.

DISCUSSION AND CONCLUSIONS

The results provided strong evidence that the three independent variables were in fact manipulated as intended. Therefore we feel justified in discussing the affectivity results in terms of these factors.

The several demonstrations of the outcome main effect indicate that consistent outcomes are preferred to discrepant outcomes. Another way of stating this is that the positive affect generated by confirmation (i.e. the occurrence of an expected, albeit weakly, event) is a positive function of expectancy strength, whereas positive affect generated by disconfirmation is an inverse function of expectancy strength. Thus some support is given to the theoretical assumption that the stronger an expectancy is, the more motivated one is to confirm and maintain it, and the weaker an expectancy is, the more motivated one is to disconfirm and to supplant it with another (presumably stronger) expectancy.

The first-order interaction between outcome and strength indicated that confirmation of expectancies (albeit weak) generates greater positive affect than disconfirmation. This result has often been demonstrated in the consistency research.

The second-order interaction among outcome, strength and value is also a rather obvious finding, indicating simply that confirmation of positive expectancies generates greater positive affect than confirmation of negative expectancies, while disconfirmation of negative expectancies generates greater positive affect than disconfirmation of positive expectancies.

The anomalous value main effect in the warm–cold responses indicated that subjects holding positive expectancies reported feeling warmer (more positive) than subjects holding negative expectancies. This is understandable, but the significant outcome × value interaction indicated that outcomes consistent with negative expectancies (strong and weak) and outcomes discrepant from positive expectancies generated stronger feelings of warmth than outcomes consistent with positive expectancies and outcomes discrepant from negative expectancies. This result is neither predicted nor explained by the theory.

The close relationship between the theoretically derived facet ordering and the empirical ordering of the eight experimental conditions provides strong evidence that the outcome variable was the most important in determining affectivity, the strength variable was next in importance, and value was least important. The principal components notion (Foa, 1965) leads one to predict that the effects of these three variables would be in a ratio of 1:2:4. Examination of the significant F ratios in the total positive affect data indicates that the obtained ratio approximated 1:2·5:3·9. Obviously this ratio depends upon the relative levels of the three variables, and since they were determined intuitively in the present study, future research should concern itself with a systematic analysis of the relevant parameters.

The only deviation from the predicted order—in the condition involving the occurrence of an unexpected positive event—is interpreted as indicating a 'pleasant surprise'. In this specific condition, the predicted negative cognitive affect generated by the confirmation of a weak expectancy was apparently masked by the positive hedonic affect generated by being allowed to taste good foods.

Consistency theories (e.g. Aronson & Carlsmith, 1962; Aronson *et al.*, 1963; Gerard, 1967) predict that persons prefer confirmation, and discrepancy theories (e.g. Berlyne, 1960; Fiske & Maddi, 1961; Fowler, 1965) predict that persons prefer disconfirmation. This study supports the conclusion that one prefers confirmation of strong expectancies and disconfirmation of weak expectancies, and therefore that consistency theoretical predictions may be restricted to persons holding strong expectancies, and discrepancy theoretical predictions may be restricted to persons holding weak expectancies.

This study is based upon a dissertation submitted to the Graduate Faculty of the University of Missouri (Columbia) in partial fulfilment of the requirements for the Doctor of Philosophy. The guiding comments and suggestions of Drs Bruce J. Biddle and Uriel G. Foa are especially acknowledged.

REFERENCES

Aronson, E. & Carlsmith, J. M. (1962). Performance expectancy as a determinant of actual performance. *J. abnorm. soc. Psychol.* **65**, 178–182.

Aronson, E., Carlsmith, J. M. & Darley, J. (1963). The effect of expectancy on volunteering for an unplanned experience. *J. abnorm. soc. Psychol.* **66**, 220–224.

Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity.* New York: McGraw-Hill.

Bramel, D. (1968). Dissonance, expectation, and the self. In R. P. Abelson *et al.* (eds.), *Theories of Cognitive Consistency: A Sourcebook.* Chicago: Rand McNally.

Day, C. R. (1961). Some consequences of increased reward following establishment of an output-reward expectation level. (Unpublished M.A. thesis, Duke University.)

Festinger, L., Riecken, H. W. & Schachter, S. (1965). *When Prophecy Fails.* Minneapolis, Minn.: University of Minnesota Press.

Fiske, D. W. & Maddi, S. R. (eds.) (1961). *Functions of Varied Experience.* Homewood, Ill.: Dorsey Press.

Foa, U. G. (1965). New developments in facet design and analysis. *Psychol. Rev.* **72**, 262–274.

Fowler, H. (1965). *Curiosity and Exploratory Behavior.* New York: Macmillan.

Gerard, H. B. (1967). Compliance, expectation of reward, and opinion change. *J. Person. soc. Psychol.* **6**, 360–364.

Manuscript received 2 June 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 234–240
Printed in Great Britain

Persuasiveness of Confidence Expressed via Language and Body Language

By CATHA MASLOW, KATHRYN YOSELSON AND HARVEY LONDON

Brandeis University

In previous studies (Johnson & London, 1968; London *et al.*, 1970*a*) expressed confidence has been identified as a new 'message' variable causing persuasion. However, the variable has been extracted from naturalistic discussion in a dyad. In order to determine further its efficacy, we manipulated expressed confidence in two studies. In Study 1 expressed confidence was manipulated via language. In Study 2 expressed confidence was manipulated via body language. The results confirm earlier findings and indicate that a 'channel' notion is required to understand the expression of confidence.

A new method of studying persuasion, the 'jury method', has recently been developed (London *et al.*, 1970*b*, *c*). This technique allows study of the natural behaviour of the persuader in the context of two-person persuasive interaction. No method has heretofore been available which permits random assignment of subjects to opposing points of view with respect to an issue and subsequent observation of subjects as they attempt to resolve their disagreement.

Studies using the jury method have presented evidence that it is the expression of confidence which causes the persuader to change the attitude of the persuadee (Johnson & London, 1968; London *et al.*, 1970*a*). The 'expressed confidence' notion may be distinguished from other, perhaps similar, variables such as credibility and self-esteem (McGuire, 1969*a*, pp. 182, 250) found to affect attitude change. Using McGuire's terminological schema (1969*a*, p. 172), credibility is a *source* variable and self-esteem is a *receiver* variable, while expressed confidence is a *message* variable.*

To be specific with respect to the 'expressed confidence' concept, the words uttered by subjects discussing an issue and attempting to resolve their disagreement may be conceived as expressing confidence (in terms of statements conveying Confidence in Self (C_s) and Doubt in Other (D_o)) and doubt (in terms of statements conveying Doubt in Self (D_s) and Confidence in Other (C_o)). In the course of discussion both person A and person B will utter a number of statements relevant to the expression of confidence and doubt and a number of statements that are irrelevant. From the relevant statements uttered by each subject one may derive a 'vector of persuasion' which, being directional, necessarily represents *both* the persuasive force exerted on the other *and*, simultaneously, the resistance to persuasion of the self. To compute the vector one performs a content analysis (Yoselson

---

* The question may arise whether expressed confidence should be considered a 'message style' or a 'message discrepancy' variable (cf. McGuire, 1969*a*, pp. 207, 217). We prefer at this stage in our research to consider expressed confidence a message style variable since such a categorization seems relatively face-valid and theory-free. To consider expressed confidence a message discrepancy variable implies more than we yet know about the mechanism by which expressed confidence persuades.

et al., 1970) of a discussion and, for each subject, simply subtracts from the number of statements indicating confidence the number of statements indicating doubt. That is, vector of persuasion = confidence − doubt = $(C_s + D_o) - (D_s + C_o)$.

The present report presents two studies which build upon the research we have outlined. The studies cited above have extracted linguistically expressed confidence from naturalistic discussion in a dyad. The inference that expressed confidence *caused* persuasion was made possible by a process of ruling out alternative explanations. This approach to causal analysis has been developed or endorsed by methodologists in a number of disciplines (Blalock, 1964; Campbell & Stanley, 1963; Li, 1955; McGuire, 1967, 1969*b*; Simon, 1957); nevertheless, any variable whose causal status has been inferred in this way should perhaps also be investigated by more conventional means. For this reason, in Study 1 we manipulate expressed confidence under stringently controlled conditions. If the amount of confidence expressed within an argument is varied while all other elements of the argument are held constant, and if a confidently expressed argument is shown more persuasive than one less confidently expressed, then the expressed confidence notion gains greater credibility.

As mentioned above, the expression of confidence has thus far been measured via language. However, it is clear that confidence and doubt may be expressed in a variety of ways and that a linguistic measure indexes only one among a number of possible modes.

Therefore, in Study 2 we explore a second potential 'channel' (McGuire, 1969*a*, p. 172) by manipulating confidence kinesically, i.e. via body language. Although body-language variables have been increasingly researched in recent years (for recent reviews, see Duncan, 1969; Mehrabian, 1969), few studies (cf. Mehrabian & Williams, 1969) have directly investigated the persuasiveness of kinesic cues. If the amount of confidence expressed kinesically is varied while other elements of an argument are held constant, and if the confidently expressed argument is more persuasive, then we will have demonstrated a second channel by means of which expressed confidence can operate.

STUDY 1

Method

Subjects

The subjects were 52 undergraduate women. They were members of a social club and participated in the study immediately prior to a club meeting.

Procedure

Subjects took part individually in a 'jury study'. Subjects were administered booklets containing the facts of a law case, instructions to the jury, an argument purported to be that of 'a student interviewed in a previous study', verdict sheets and, finally, a page on which the subjects were asked to describe their impression of the student whose argument they had read. Subjects read the materials in the booklet and registered their verdicts and impressions.

Communications

Booklets contained the same fact summary and judge's instructions used previously (London *et al.*, 1970*c*). Written especially for the present study was the 'student's argument' in favour of finding the defendant, Brooks Airlines, not liable.

Student's argument. There were two versions of the argument for finding the defendant not liable. Both versions made the same points in favour of the defendant. However, the argument was structured so that there were 'slots' into which expressions of confidence or doubt could be placed. These expressions were drawn from the scoring manual mentioned above (Yoselson *et al.*, 1970). One half of the subjects, randomly chosen, received the argument containing the confidence expressions; the other half received the argument containing the doubt expressions. The tone of the arguments so constructed did not differ from the tone of arguments used by 'real' persuaders and persuadees (cf. Johnson & London, 1968; London *et al.*, 1970*a*).

The argument is reproduced below. It was written colloquially, and was presented to the subjects as if it were a transcript of an interview with a student. Pairs of expressions in parentheses are the aforementioned expressions of confidence and doubt. For each pair, the first would be placed closer to the 'extreme confidence' end of a continuum running from confidence to doubt. A reading of the two separate versions of the argument will reveal that while in both the student reaches the same conclusions for the same reasons on each point, the expressions in the doubtful version suggest that the student is wavering before he reaches each conclusion.

'(Obviously, I don't know) my first (concern, conflict) is deciding what to base liability on. I mean, we were given only the facts, so I have to (reason, wonder) what the law must say about this. (I believe, I'm not positive) that causation is the best basis for judging whether someone should pay damages. Looking at it (logically, all kinds of ways) the (clarity, ambiguity) of the meaning of causation is (evident, confusing) to me. I would say with (assurance, hesitation) that the cause of any damage is that without which the damage would not have occurred.

'Another (point, question) I'm quite (sure, unsure) about is the mechanic's negligence. (I feel, I'm not sure) it's irrelevant. I would say with some (confidence, doubt) that the case should be based solely on causation . . . [pause].

'If Brooks Airline Company's liability depends on causation, this case rests on whether or not the damage would have occurred even if the plane hadn't crashed. (I am certain, My problem is) that you can't know whether the natural fire would have caused Ames Lumber Company's loss whether or not the plane had crashed. Since either fire may have caused the damage (it's obvious to me, I'm leaning both ways), (you can't assume that, I don't know if) the plane fire caused the damage. (I'm satisfied, I'm unsure) that there's only one fair verdict. In fact, my verdict (is determined, could go either way). I'll have to base it on this, that (it's clear, I admit) that you can't assume that Brooks Airline Company is the cause of the damage. Therefore, I (maintain, grant) that we can't hold them liable for Ames Lumber Company's loss'.

Measurement of persuasion

Persuasion was assessed by the extent of the subject's agreement with the hypothetical student that Brooks Airlines, the defendant, should not be held liable for the plaintiff's damage. Subjects were asked to indicate on the verdict sheets immediately following the substantive material (*a*) percentage *confidence* (from 0 to 100) that Brooks should not be held liable; (*b*) percentage *doubt* (also from 0 to 100) that Brooks should not be held liable.

RESULTS

Effectiveness of manipulation

To test whether expressed confidence was effectively manipulated, Thurstone's paired comparison technique (Edwards, 1957) was utilized. Five female college students, paid volunteers, were recruited from Boston University's summer school introductory psychology class. These students were tested simultaneously. They were given booklets containing instructions and both versions of the student's argument. In three of the booklets, the confidently expressed argument was placed first; in two of the booklets, the doubtfully expressed argument was placed first. The girls were assigned booklets at random and were asked to indicate on a rating sheet which of the versions of the argument expressed more confidence in the points presented. All five girls rated the version postulated to express more confidence as the more confident. The P value associated with this result is 0·031.* This P value indicates that the expression of confidence was effectively manipulated.

Effect of linguistically expressed confidence on persuasion

The hypothesis advanced was that subjects receiving the confidently expressed argument would be significantly more persuaded than subjects receiving the doubtfully expressed argument.

To obtain an overall index of confidence that Brooks should not be held liable, we simply subtracted a subject's percentage of doubt from his percentage of confidence.† Subjects receiving the confident argument had a mean confidence index of 44·0 that Brooks should not be held liable. Subjects receiving the doubtful argument had a mean confidence index of 10·3. A t test of the difference between these means yielded $P < 0·05$. Subjects receiving the confident argument were significantly more persuaded that Brooks Airlines should not be held liable.

STUDY 2

Method

Subjects

The subjects were 69 undergraduate females, all of whom were paid volunteers.

Manipulation of kinesically expressed confidence

A professional actor was hired to play a law student presenting his views about a case. The actor, dressed conservatively and seated at an executive-sized desk cluttered with papers and books, was filmed using a Sony videotape camera and a Sony videocorder. The camera rested on a table, about 7 ft. away from the actor.

Before filming, the actor tape-recorded a 'discussion' of the case. The discussion consisted of the argument used in Study 1 but written 'neutrally', i.e. with confidence and doubt words deleted and other words inserted appropriately. The tape-recorded discussion allowed control of language (words) and para-language (tone of voice, rate of speech, etc.) across conditions insofar as the discussion was 'dubbed' into all three videotapes used in this study. The tape discussion also facilitated control of the length of videotapes at 1 min. 10 sec.

Three separate videotapes were prepared. For one, the actor was asked to 'discuss' the case in a confident manner, for the second, in a doubtful manner, and for the third, 'neutrally', i.e. midway between confidence and doubt. In none of the three videotape sessions did the actor actually speak. Rather, he coordinated his mouth movements to coincide with his words on the previously recorded sound tape, which was played during filming. Later, pretesting indicated that the subjects were unaware that the actor was simply mouthing the words he appeared to be uttering.

Procedure

The experimenter scheduled the subjects for experimental sessions by circulating sign-up sheets. In all, six sessions were conducted, two per experimental condition. Twenty-four subjects were run in the confidence condition, 20 were run in the neutral condition, and 25 were run in the doubt condition.

At each session, the subjects were seated and asked to make sure they had a full view of a Sony television monitor which was placed on a high desk in the front of the room. The experimenter then explained that the point of the study was to help evaluate the ability of

---

[*] All P values reported in this paper are two-tailed.

[†] This mode of combining response scales to yield a more sensitive index is suggested in Schachter & Singer (1962, p. 389). It should be noted that the two components of the index each yield results completely consistent with those obtained by the use of the index.

law students to work in legal clinics. The subjects were asked to play the role of jury members in a law case. They were to read the facts of a law case and judge's instructions to the jury, and then to observe a videotape with comments on the case by a high-ranking law student. They were told finally that their individual verdicts would be solicited after they had viewed the tape.

While subjects were reading the background materials, precisely the same as those used in Study 1, the experimenter chose a slip of paper from an envelope in order to determine which of the three tapes would be shown.

After the film the subjects indicated their verdicts, thus providing the experimenter with a measure of persuasion, and filled out a post-experimental questionnaire.

Finally, the experimenter collected all materials, gave the subjects further information about the study, thanked the subjects for their participation, and paid them.

RESULTS

Effectiveness of manipulation

To test whether expressed confidence was effectively manipulated, Thurstone's paired comparison technique was again utilized. Ten female high school students, paid volunteers, were recruited. These students were tested one at a time and were asked to indicate on a rating sheet for all possible pairs of the three videotapes, which tape of each pair expressed more confidence in the points presented. The order of pairs and the order of tapes within each pair were randomized using an appropriate procedure.

The resulting data were analysed to yield the following scale values: doubtful tape, 0·00; neutral tape, 2·92; confident tape, 3·27. These values are in line with expectations. Further Mosteller's (1951) test of significance for paired comparisons indicated ($0.50 < P < 0.70$) that the assumptions involved in using the method of paired comparisons are tenable for the given data.

Effect of expressed confidence on persuasion

As in Study 1, persuasion was assessed via the extent of subject's agreement with the speaker, i.e. subject's overall confidence that Brooks Airlines should not be held liable. From the data on the subject's verdict sheet, confidence indices were constructed by subtracting the subject's percentage confidence that Brooks should be held liable from her percentage confidence that Brooks should not be held liable. Mean indices ran from a low of 7 in the doubt condition to a high of 46·9 in the confidence condition. Analysis of variance indicated that confidence indices differed as a function of experimental condition ($F = 3.66$; d.f. = 2, 66; $P < 0.05$). Fitting appropriate coefficients to determine the significance of the linear component of the between-condition sum of squares yielded $F = 7.3$, d.f. = 1, 66, $P < 0.01$. Subjects were increasingly persuaded as a function of increasingly kinesically expressed confidence.

DISCUSSION

The results support both hypotheses and allow us to begin building a 'structure' using McGuire's (1969*a*, p. 172) terminology: a *source* is increasingly persuasive as his *message* increases in confidence whether expressed over linguistic or kinesic *channels*. Putting the data in this way immediately suggests a third possible channel for expressing confidence—the paralinguistic. Persumably, the relation between paralinguistically expressed confidence and persuasion would parallel the findings of the present report.

However, knowledge of within-channel relationships does not necessarily allow us to determine what will occur when expressed confidence is varied in two or more channels simultaneously. The most obvious prediction is that effects will simply summate, but there are other possibilities. For instance, a person who expresses much confidence over one channel and much doubt over another might appear odd and thus be less persuasive than another person who expresses intermediate confidence over both channels, even though both people express the same 'net' confidence.

In any event, the channel notion does have validity for distinguishing between the various modes by means of which confidence may be expressed. At the same time, Study 2 highlights a peculiarity of the source/message distinction. We conceptualized our manipulations in both Studies 1 and 2 as of attributes of the message. Yet, in Study 2 at least, we could hardly have manipulated kinesically expressed confidence without ourselves operationally manipulating the source. Perhaps what this shows is that certain variables, e.g. expressed confidence, are message attributes only for certain theoretical purposes and that, for other purposes, they must be conceived to be attributes of the source as well.

As a final point we note that the results of Study 2 raise the question of differences in the actor's behaviour across conditions. We do not provide an analysis of differences here, since the focus of Study 2 was simply to establish the existence of a non-verbal channel for the expression of confidence; to go into the details of the kinesic cues involved would disproportionately weight the paper in this respect. Such an analysis, however, is an obvious next step.

This research was supported by N.I.M.H. grant 2TO2MH-07701-06 and by N.I.H. grant 7044-04. We thank Fred Grandy who played the law student in Study 2 and Michael Grinberg who aided in data analysis.

REFERENCES

Blalock, H. (1964). *Causal Inferences in Nonexperimental Research.* Chapel Hill: University of North Carolina Press.

Campbell, D. T. & Stanley, J. C. (1963). *Experimental and Quasi-experimental Designs for Research.* Chicago: Rand McNally.

Duncan, S. (1969). Nonverbal communication. *Psychol. Bull.* **72**, 118–137.

Edwards, A. L. (1957). *Techniques of Attitude Scale Construction.* New York: Appleton-Century-Crofts.

Johnson, R. & London, H. (1968). Perceived confidence, expressed confidence, and resolution of disagreement. (Paper read to American Sociological Association, Boston, Mass.)

Li, C. C. (1955). *Population Genetics.* Chicago: University of Chicago Press.

London, H., Meldman, P. J. & Lanckton, A. V. C. (1970*a*). The jury method: how the persuader persuades. *Publ. Opin. Q.* **34**, 171–183.

London, H., Meldman, P. J. & Lanckton, A. V. C. (1970*b*). The jury method: some correlates of persuading. *Hum. Relat.* **23**, 115–121.

London, H., Meldman, P. J. & Lanckton, A. V. C. (1970*c*). The jury method. (National Auxiliary Publications Service.)

McGuire, W. J. (1967). Some impending reorientations in social psychology: some thoughts provoked by Kenneth Ring. *J. exp. soc. Psychol.* **3**, 124–139.

McGUIRE, W. J. (1969*a*). The nature of attitudes and attitude change. In G. Lindzey & E. Aronson (eds.), *The Handbook of Social Psychology*, 2nd ed., vol. 3. Reading, Mass.: Addison-Wesley.

McGUIRE, W. J. (1969*b*). Theory-oriented research in natural settings: the best of both worlds for social psychology. In M. Sherif (ed.), *Interdisciplinary Relationships in the Social Sciences*. Chicago: Aldine.

MEHRABIAN, A. (1969). Significance of posture and position in the communication of attitude and status relationships. *Psychol. Bull.* **71**, 359–372.

MEHRABIAN, A. & WILLIAMS, M. (1969). Nonverbal concomitants of perceived and intended persuasiveness. *J. Person. soc. Psychol.* **13**, 37–58.

MOSTELLER, F. (1951). Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed. *Psychometrika* **16**, 207–218.

SCHACHTER, S. & SINGER, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychol. Rev.* **69**, 379–399.

SIMON, H. A. (1957). *Models of Man*. New York: Wiley.

YOSELSON, K., MELDMAN, P. J. & LONDON, H. (1970). A scoring manual of confidence and doubt. (National Auxiliary Publications Service.)

Manuscript received 8 June 1970

Revised manuscript received 11 December 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 241–252
With 6 *text-figures*
Printed in Great Britain

Factors Affecting Word Selection by Schizophrenic Patients

By GABRIELLE MARCHBANKS* and MOYRA WILLIAMS

Fulbourn and the United Cambridge Hospitals

Experiments were done to discover some factors relating to word selection by schizophrenic patients. The comprehension of individual words was examined by presenting the subjects with a picture of an object and a number of words containing the name of the object, from which the object name had to be selected.

Schizophrenics were distinguished from other groups of patients, from children and from normal controls by accepting associations to the object as well as the object name. The frequency of acceptance was directly related to the proximity of the association.

It was also found that schizophrenics accepted associated words more often if the stimuli were presented in an open-ended or single-choice manner than if they were presented in a multiple-choice form. The presentation of other words, with which comparisons could be made, further reduced acceptances.

Schizophrenic patients took longer to reject words than control subjects, and when under time pressure increased their acceptance of associated words.

It is suggested that although their word store is organized in a normal manner, schizophrenic patients may have greater difficulty than normal people in recognizing the boundaries between associated words, and that this may be the basis for many aspects of their communication difficulties.

Defects of communication are a basic sympton of schizophrenic illness. In past years this has often been attributed to difficulties of expression, but recently it has been suggested that a disorder of comprehension may be involved, this disorder being variously ascribed to impairment in the mechanism of selective attention and filtering (Chapman & McGhie, 1963; Ludwig *et al.*, 1963; McGhie & Chapman, 1961; McGhie *et al.*, 1965), to cortical over-arousal (Venables, 1969), to a loosening of conceptual boundaries (Bannister, 1960), to a disturbance in the breadth of associations (Craig, 1967; Gathercole, 1965) and to overinclusiveness (Epstein, 1953; Payne & Friedlander, 1962).

The experiments to be described here were designed to examine some of the variables affecting the extent to which schizophrenic patients differentiate between words as the names for objects, in the hope that a study of this function might throw light on the mechanisms which are disturbed by schizophrenic illness.

The schizophrenic patients who acted as subjects in the experiments were all in-patients at Fulbourn Hospital, Cambridge, who had been diagnosed by the consultant in charge. For every experiment, care was taken to ensure that the subjects contained equal numbers of men and women, and equal numbers of chronic (those whose illness was of at least three years' duration) and acute subjects (patients in their first or second admission to hospital). In the original analysis of each experiment, distinction was made between the paranoid and non-paranoid, and between the thought-disordered and non-thought-disordered patients. However,

---

* On a grant from the Mental Health Research Fund.

as the differences between chronic and acute, paranoid and non-paranoid, thought-disordered and non-thought-disordered were not always significant, the performance of all the schizophrenic patients will be presented here together.

In order to eliminate as many variables as possible on the expressive side of language a basic technique was used in which the subject was presented with a picture of a single object and one or more words, with the instruction to match the picture with its correct name. Analysis was then made of the names accepted and rejected.

EXPERIMENT 1

This was a pilot experiment aimed at comparing schizophrenic patients with other members of the community in the acceptance of names for objects.

Procedure. The task was prepared from 12 object nouns and 12 pictures of these objects. Each object had one unequivocal name in the English language. Nine of the names were equally divided into groups of one-, two- and three-syllable words, and into three groups varying in word frequency according to the Thorndike & Lorge (1944) word count. In addition to these nine object nouns, one composite noun of three syllables and its two component parts were included in the task (see Table 1).

Attached to each of these object nouns, eight alternative words were selected to represent phonetic similarity (two words began with the same syllable as the correct object noun and two words ended with the same syllable as the correct object noun), associational relationship (two words were related in an associative manner to the object noun) and irrelevance (two words were not related to the object noun).

These eight alternative words, together with the object noun, made up one item in the 12-item task. The words within each item were arranged in random order and the items were also arranged in random order. The words were recorded in a standard English voice on tape at the rate of one per 5 sec.

The items were then presented to the subjects under two conditions referred to as the multiple- and single-choice conditions. The task was divided in half and the halves were systematically alternated from one subject to another. Under the multiple-choice condition the subject was told, 'You will hear a list of words on the tape. When you hear the name of this picture or what this is a picture of, please say "yes". For example, if I give you a picture of a dog and the words "cat" and "log" come up on the tape, you should remain silent. But when you hear the word "dog" which is the name of this picture you will say "yes".'

Under the single-choice condition the subject was instructed to respond 'yes' or 'no' to each name.

Each time he responded positively to a word which was not the correct name he was asked if the word he responded to was the *name* of the object. If the subject once again stated that it was the name, a note indicating an error was made. Subjects often explained why they chose to accept an incorrect word and the reasons were noted.

Finally, each subject was shown the 12 pictures once again and asked to name them one at a time.

Subjects. The task was administered to 60 schizophrenic patients and to the following control groups. *Normal adults:* 10 patients (five men, five women, mean age 29·2, range 16–63) in the orthopaedic ward at Addenbrooke's Hospital, Cambridge, without any psychiatric history. *Children:* 10 five-year-old children (five boys, five girls). *Dysphasics:* 10 patients (four men, six women, mean age 52·6, range 24–71) in the neurological wards at Addenbrooke's Hospiral, Cambridge, demonstrating expressive dysphasia but no comprehension defect. *Senile dements:* 10 patients (five men, five women, mean age 77·4, range 63–89) resident in Fulbourn Hospital, Cambridge. *Non-psychotic psychiatric patients:* 10 patients (five men, five women, mean age 33·0, range 19–49) from the admission wards of Fulbourn Hospital, Cambridge, diagnosed as suffering from personality disorders or neurotic illness.

Results

All results were evaluated by the Mann–Whitney *U* test (two-tailed) or the Wilcoxon test (two-tailed).

The main results can be summarized as follows.

1. Schizophrenic subjects made significantly more errors by accepting incorrect names than non-psychiatric patients ($P < 0{\cdot}002$) and than normal adults ($P < 0{\cdot}002$). The demented patients made significantly more errors than the schizophrenics ($P < 0{\cdot}002$).

2. Significantly more incorrect names were accepted by schizophrenic subjects under the single-choice than under the multiple-choice condition ($P < 0{\cdot}01$).

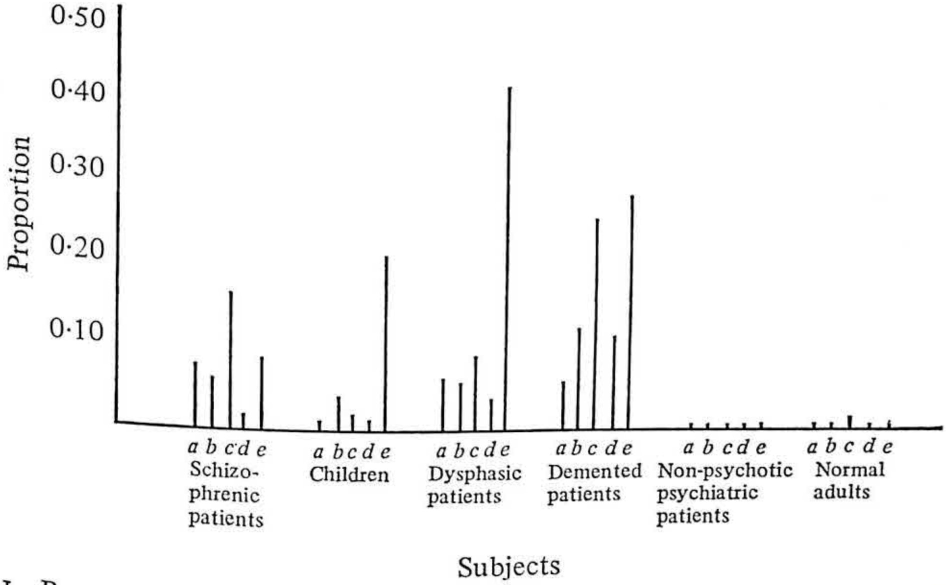


Fig. 1. Proportions of types of errors made by schizophrenic and control subjects. (*a*) Phonetically similar at the beginning of the word. (*b*) Phonetically similar at the end of the word. (*c*) Related in meaning. (*d*) Irrelevant. (*e*) Rejection of the correct word.

3. Significantly more errors were made on long stimulus words than on short words by the schizophrenic patients ($P < 0{\cdot}01$), and by children ($P < 0{\cdot}01$) but word frequency only affected children.

4. Schizophrenic patients did not differ significantly from normals and non-psychiatric patients in naming the pictures.

5. The errors of acceptance made by schizophrenic patients differed from those made by any other group of subjects in consisting predominantly of meaning-related words (see Fig. 1).

6. Children and dysphasic subjects made significantly more ($P < 0{\cdot}01$) rejection errors than any other kind.

Examination of the reasons given by schizophrenic subjects for accepting the wrong words showed that some associational relationship between the correct word and the incorrectly accepted one had always been found. For example, asked if the word 'infant' was the name of a picture of an elephant, a schizophrenic subject accepted it because it was an 'infant elephant'. A few other examples of remote associations given as reasons for accepting phonetically similar words were 'finish' for 'finger' 'because it ends', 'sparrow' for 'arrow' 'because you can shoot them', 'triumph' for 'elephant', 'because when he blows his horn it is triumphant'.

The only other group of subjects to make a large number of acceptance errors was the demented group, but the explanations offered by them usually revealed visual misperception of the pictures. For example, 'nail' was accepted for 'dice' 'because it looks like nails in a box', 'rainbow' for 'window' 'because it is an open picture of a rainbow', 'marrow' for 'arrow' 'because it looks like one'. The demented patients, who also differed significantly from the schizophrenics in their naming ability, showed evidence of visual defects in this latter task.

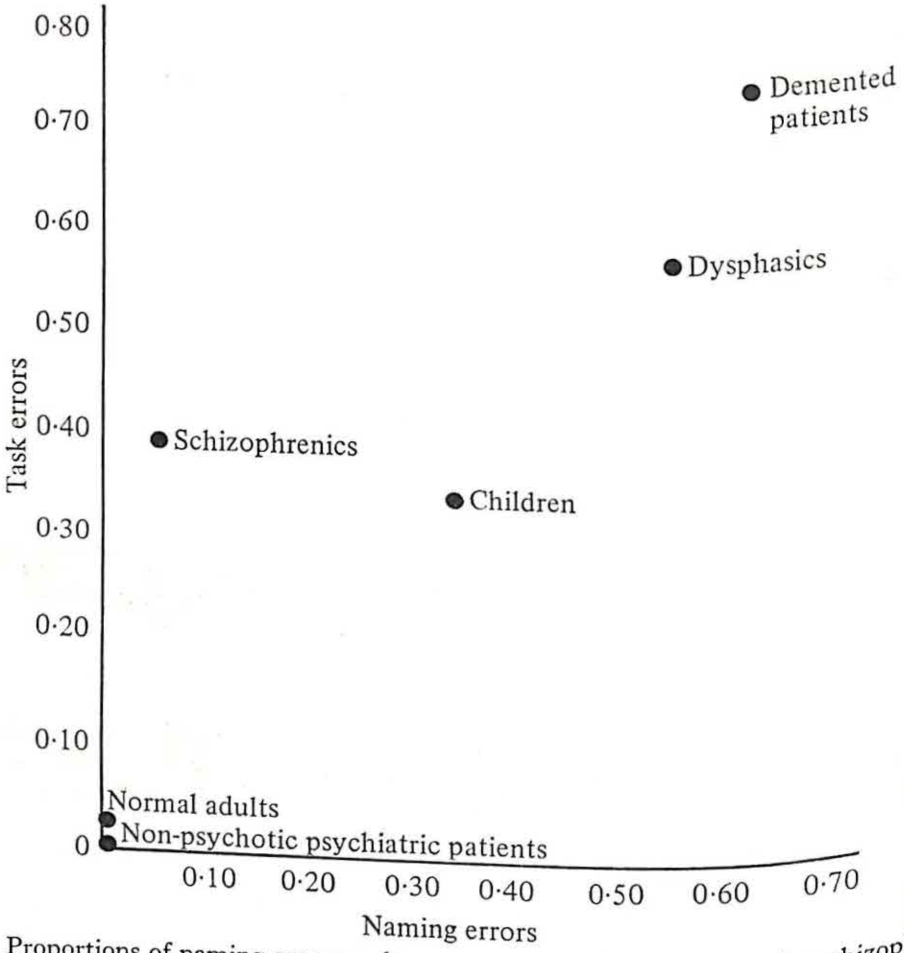


Fig. 2. Proportions of naming errors and proportions of task errors made by schizophrenic and control subjects in Expt. 1.

The most frequent error made by five-year-old children and by dysphasics was rejection of the correct name. Acceptance of an incorrect name was rare in both groups, and when this occurred reasons could rarely be elicited. The names rejected by children were almost entirely the rare and long words.

A direct relationship might have been expected between errors in name-selection and errors in naming. This is indeed the case for all groups of subjects tested except the schizophrenics (see Fig. 2).

However, the manner in which the pictures were named by the schizophrenic patients deserves comment. Although the correct name was always given, there was a tendency for it to be elaborated with inessential details, e.g. 'An elephant with ears', 'A happy elephant', 'An elephant at the zoo'. These elaborations often included words which were accepted as the names themselves when offered as alternatives, and suggest that such elaboration springs from their difficulty in inhibiting associated words.

The findings regarding the schizophrenic patients were later confirmed with a different set of words and pictures presented to different patients.

Examination of the words accepted so readily as names by the schizophrenic patients in Expt. 1 shows that the majority could be regarded as associations to the object names. However, no attempt had been made in that experiment to control the words offered for associational strength or proximity. The next experiment was undertaken to examine the relationship between associational proximity and the likelihood of a word being accepted by a schizophrenic subject.

EXPERIMENT 2

Eighty-five normal subjects (non-medical hospital staff, university students, and patients from the orthopaedic wards) were asked to write down their first associations to each of the object names used in Expt. 1 (12 items) and the confirmatory study (six items) (see Table 1). The first, second, third and fourth most popular association to each object name, together with four rare associations (i.e. associations given by not more than one of the 85 subjects), were then combined with the correct name as the alternatives in a new task of 18 items which was administered like Expt. 1 to 20 schizophrenic patients, 10 non-psychotic psychiatric patients and 10 normal subjects. The results are shown in Fig. 3, from which it will be seen that there is a close relationship between associational proximity and the number of acceptance-errors made by the schizophrenics.

Table 1. *Names of objects used as stimuli in Expt. 2*

*Window	*Dice	*Gondola
*Newspaper	*Elephant	*Fingernail
*Arrow	*Nail	*Cup
Fish	Jug	*Table
*Finger	Giraffe	Acrobat
Hatchet	*Boot	Indian

* Stimuli used in Expt. 1.

In this experiment the normal and non-psychotic psychiatric subjects accepted more words than in Expt. 1, but examination showed that some of the associations offered were also synonyms for the correct object name. When these synonyms were removed, the proportion of errors made by the two control groups was reduced to zero, but that made by the schizophrenic subjects remained high.

This observation raised the question of the relationship between synonyms and associations, and the possibility that schizophrenic patients do not distinguish as clearly between the two as do normal people. This was investigated in the following experiment.

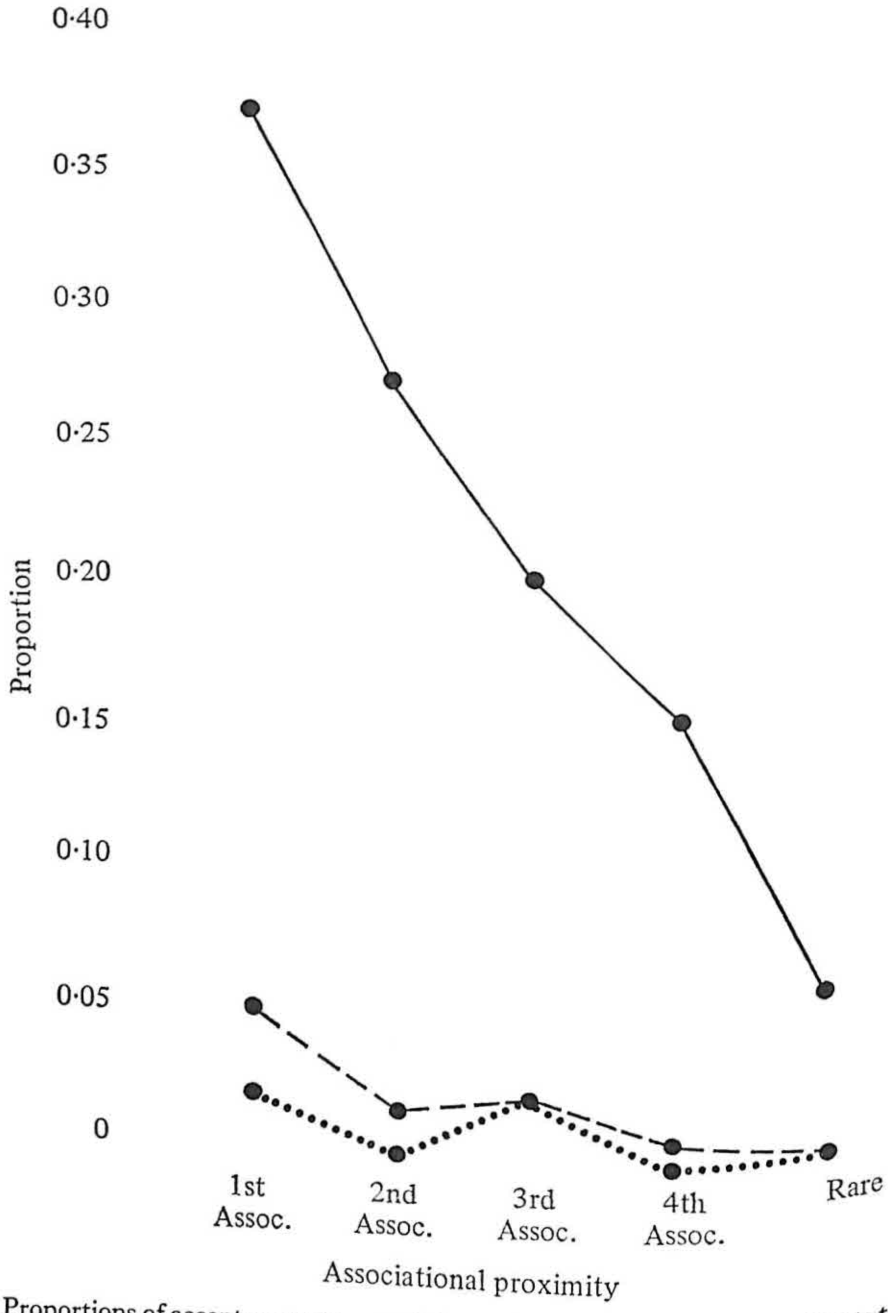


Fig. 3. Proportions of acceptance errors made by schizophrenic and control subjects to words of decreasing associational proximity to the correct object noun. ●——●, schizophrenic subjects; ●······●, non-psychotic psychiatric patients; ●------●, normal subjects.

EXPERIMENT 3

Twenty-five schizophrenic patients and 25 matched controls were asked to rate 17 words (see Table 2) comprising three key words, two synonyms to each one, two associations to each one and two irrelevant words on a Semantic Differential (Osgood *et al.*, 1957).

The scoring was based on the number of spaces differing on each dimension for each subject between each key word and its two synonyms, its two associations and the two irrelevant words. The scores were then added together, and each subject was given a total difference score for synonyms, associations and irrelevant words (see Fig. 4).

Table 2. *Words rated on the Semantic Differential in Expt.* 3

Key words	cup	hatchet	newspaper
Synonyms	mug	chopper	journal
	goblet	axe	periodical
Associations	saucer	wood	read
	tea	Indian	print
Irrelevant words	hospital, giraffe		

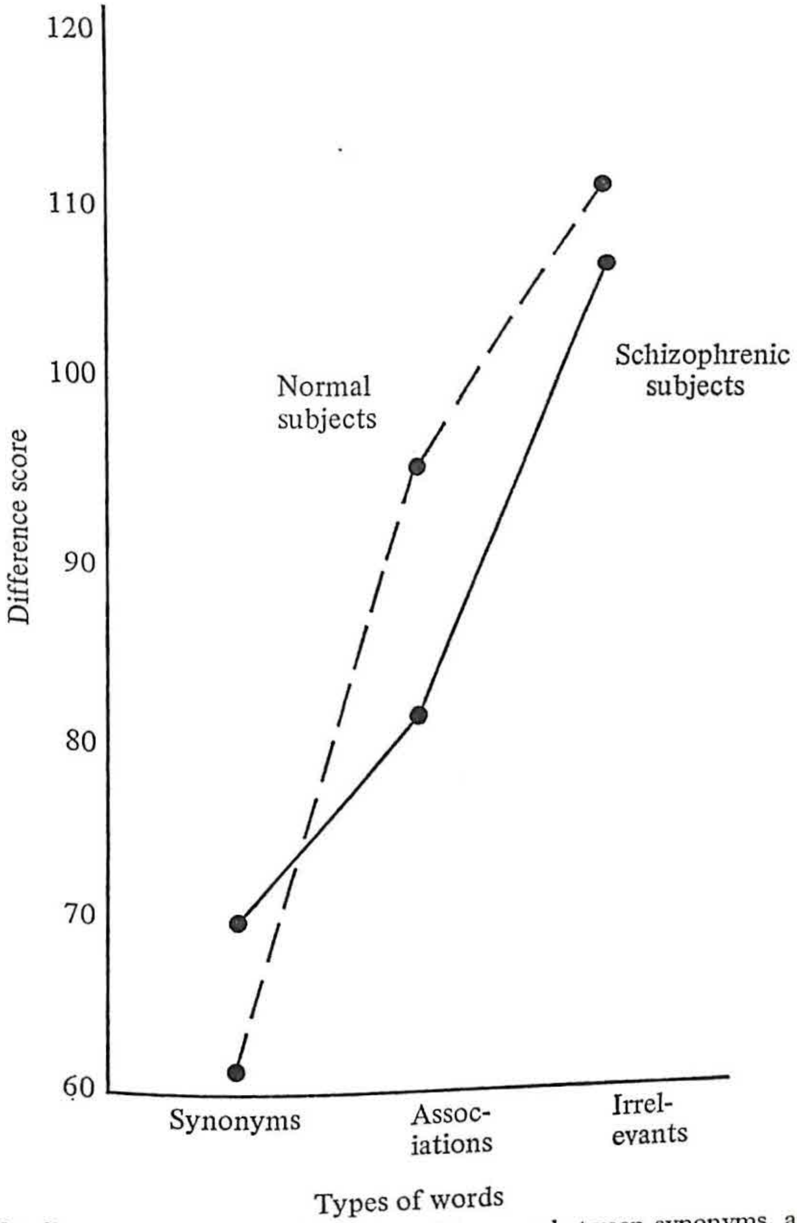


Fig. 4. Mean differences on Semantic Differential scores between synonyms, associations and irrelevant words made by schizophrenic and normal control subjects.

Both groups of subjects rated synonyms, associations and irrelevant words as significantly different from one another. However, normal subjects showed a greater difference between their ratings of synonyms and associated words than did the schizophrenics. While the schizophrenics and controls did not differ in their ratings of synonyms and irrelevant words they did differ in their ratings of the associations ($P = 0{\cdot}057$).

In Expt. 1 it was also noted that when names were offered under the single-choice condition, more acceptance errors were made by schizophrenic patients than under the multiple-choice condition. This discrepancy was confirmed in all subsequent studies.

Various hypotheses were considered to account for it. The first concerned decision-making time. Since, in the single-choice situation the subject was allowed unlimited time to make his responses to each word, whereas in the multiple-choice situation the next word was presented automatically if no response had been given within 5 sec., there was the possibility that many potential acceptances were eliminated in the multiple-choice situation by the experimenter interrupting the subject before he had signified his acceptance of a name.

The fact that schizophrenic patients took longer to reject inappropriate names (though not to accept appropriate ones) than the control subjects was verified by measuring response-time with a stop-watch. On this basis it was predicted that if placed under time pressure fewer acceptance errors will be made by schizophrenics in the single-choice situations than if they are given unlimited time as before. Expt. 4 was carried out to test this hypothesis.

EXPERIMENT 4

Twenty schizophrenic patients were presented with the same pictures and words as in Expt. 1. In half the items on the task, the subject was given unlimited time to make his responses (slow condition), but in the other half he was instructed that he would only be given 1 sec. to respond to each word (fast condition).

Contrary to prediction, significantly more errors were made in the fast condition than in the slow ($P = 0{\cdot}01$). Thus time pressure, far from decreasing the tendency to accept associations, seems instead to increase it.

The second hypothesis considered to account for the single/multiple choice difference concerned the effect of additional cues. In the multiple-choice condition, the subjects were provided with a number of alternatives which they could compare with one another. This process of comparison may have assisted rejection. On this basis it was predicted that the more alternatives they are given, the better the performance of schizophrenic patients will be. This hypothesis was tested in Expt. 5.

EXPERIMENT 5

Twenty schizophrenic patients were presented with the 18 pictures and words from Expt. 2, divided into six groups of three each. Each group of items was presented in one of the following conditions: (1) Single-choice condition, nine alternatives (the correct word, four close associations and four rare associations, as in Expt. 2). (2) Single-choice condition, five alternatives (the correct word, two close associations and two rare associations). (3) Single-choice condition, two alternatives (the correct word, and one associated word, alternating between a close and rare association from one two-alternative item to another). (4) Multiple-choice condition, nine alternatives (same as condition 1 above). (5) Multiple-choice condition, five alternatives (same as condition 2 above). (6) Multiple-choice condition, two alternatives (same as condition 3 above).

As before, more errors were made under the single- than under the multiple-choice condition, but as the number of alternatives increased the proportion of errors decreased under both conditions (see Fig. 5). It is to be noted that this occurred even in the single-choice condition although the subject was not able to make direct comparisons. Moreover, the subjects never knew in advance how many alternatives they would be offered.

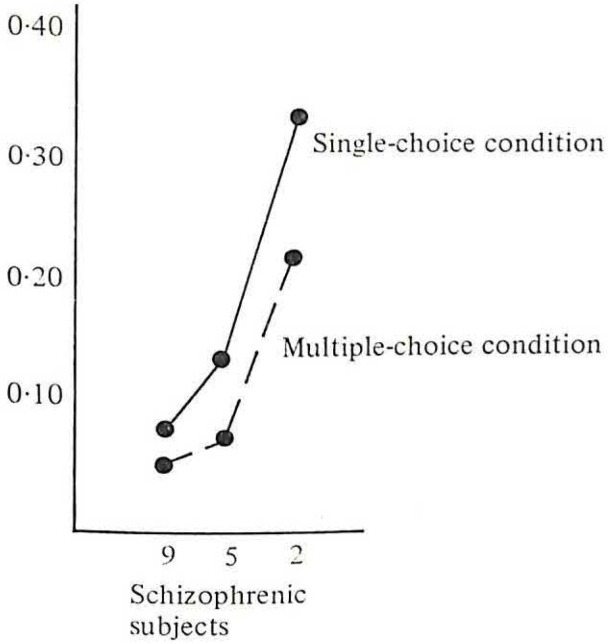


Fig. 5. Proportions of errors made by schizophrenic subjects in the multiple-choice and single-choice conditions as the number of alternatives decreases. 9 = nine alternative choices; 5 = five alternative choices; 2 = two alternative choices.

EXPERIMENT 6

Sixteen of the items and alternatives used in Expt. 2 were presented to 20 schizophrenic patients in the single-choice situation under four conditions as follows: (1) The single-choice condition, as previously administered. (2) The subject was required to read a list of irrelevant words before each item, and was then tested under the single-choice condition as before. (3) The subject was required to read the eight words associated with the correct object noun, and was then tested under the single-choice condition as before. (4) The subject was asked to read and define the meanings of each of the eight words associated with the correct object noun, and was then tested under the single-choice condition as before.

The proportion of errors under each of the conditions is shown in Fig. 6. Condition 1, which produced most errors, was significantly different from condition 2 ($P < 0·01$) and 2 was significantly different from 4 ($P < 0·05$) but not from 3.

DISCUSSION

The experiments described in this paper were designed to elucidate some of the factors involved in word-choice by patients suffering from schizophrenia. Errors of over-acceptance by schizophrenics as opposed to errors of omission by controls

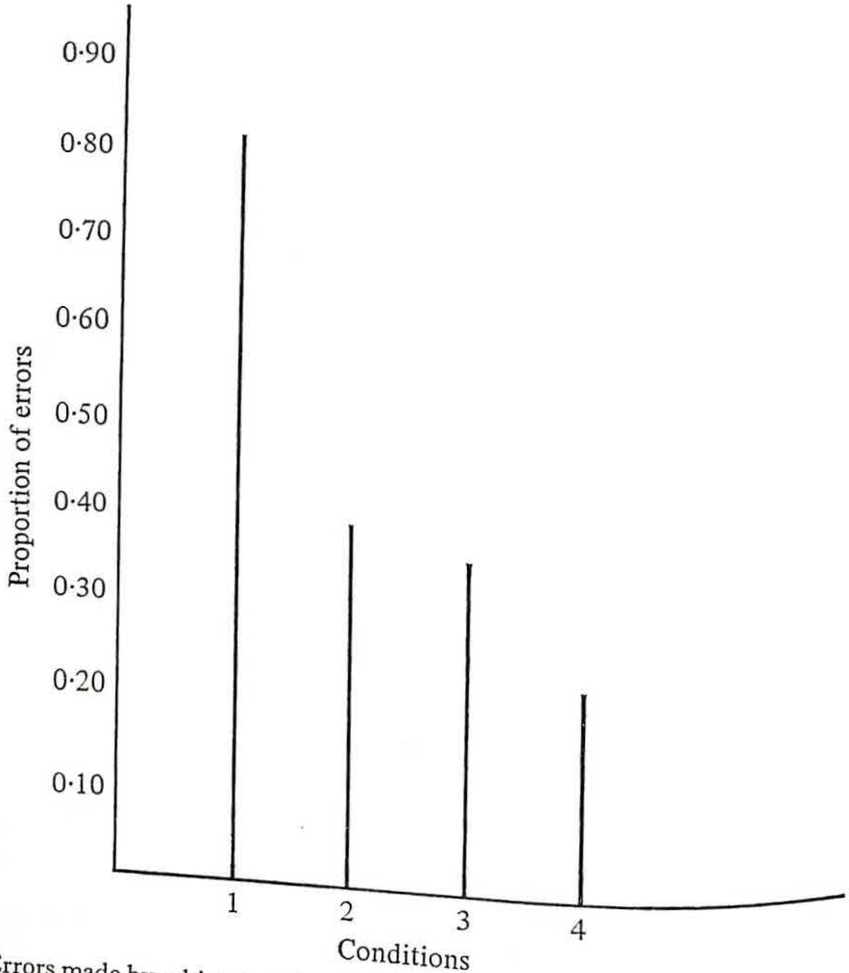


Fig. 6. Errors made by schizophrenic subjects with and without aids to boundary formation. Conditions: 1, no preparation; 2, reading list of irrelevant words; 3, reading list of associated words; 4, discussing list of associated words.

have been noted by other workers (Spence & Lair, 1964) and this observation has been made the basis of a diagnostic test by Epstein (1953) and later revised by Craig (1967). Gathercole (1965) has described schizophrenic test performance as 'an inability to restrain responses', resulting from an abnormal 'fluency of associations' generally present in schizophrenics.

This characteristic of schizophrenia has sometimes been attributed to defects in the mechanism of attention and of filtering. However, it was shown in Expts. 1 and 2 that schizophrenic patients are able to filter out or reject material which is either irrelevant or phonetically related, so that if there is an impairment in the filtering mechanism of attention, it is only with phenomena associated to the stimuli that this inability becomes evident.

It was noted in Expt. 3 that schizophrenic patients rated associations closer to one another than normal subjects on the Semantic Differential, but at the same time, Expt. 2 indicates that even with schizophrenics, associations differ in their proximity to one another. This implies that the organization of the associational repertoire is qualitatively similar for schizophrenics and for normals (substantiating the views of Cohen & Camhi, 1967; Sommer *et al.*, 1960; Johnson *et al.*, 1964), even though access to the repertoire as a whole may be different in the two groups.

To envisage how this could be accounted for, it is necessary to consider the mechanism possibly involved in word selection in general. In order that one word may be differentiated from another, each word must be enclosed by some form of conceptual boundary, based on the number of dimensions it occupied in common with or distinct from other words. Hence in order to differentiate between two words or to fit a name to an object, the dimensions of each must be scanned and compared. The final stage of acceptance or rejection must be a decision-making one, whose result is dependent on the criteria (i.e. number of dimensions in common) adopted by the judge.

The over-acceptance of words by schizophrenic patients could be due to (1) slow scanning, (2) scanning of too few dimensions, or (3) low criterion for accepting the judgement 'same'. Observations based on reaction time studies suggest that scanning is slow in schizophrenics, but even where there is no time pressure, schizophrenic patients still rate associations as more similar to one another than do normal people. The most probable explanation for the sum total of the findings reported here seems to be that as scanning does take schizophrenic patients a long time, it is usually not carried out very far, so that subjects are inclined to accept a low criterion of matching for the judgement of 'same', especially if under time pressure, as in Expt. 4.

The presentation of additional stimuli such as those reported in Expts. 5 and 6 alters the subjects' criteria for accepting 'same', encourages them to scan more systematically and so, in general terms, makes them more critical.

REFERENCES

BANNISTER, D. (1960). Conceptual structure in thought-disordered schizophrenics. *J. ment. Sci.* **106**, 1230–1249.

CHAPMAN, J. & MCGHIE, A. (1963). An approach to the psychotherapy of cognitive dysfunction in schizophrenia. *Br. J. med. Psychol.* **36**, 253–260.

COHEN, B. & CAMHI, J. (1967). Schizophrenic performance in a word-communication task. *J. abnorm. Psychol.* **72**, 240–246.

CRAIG, W. (1967). Breadth of association in psychiatric patients. *J. clin. Psychol.* **23**, 11–15.

EPSTEIN, S. (1953). Overinclusive thinking in a schizophrenic and a control group. *J. consult. Psychol.* **17**, 384–388.

GATHERCOLE, C. (1965). A note on some tests of overinclusive thinking. *Br. J. med. Psychol.* **38**, 59–62.

JOHNSON, R., WEISS, R. & ZELHART, P. (1964). Similarities and differences between normal and psychotic subjects in responses to verbal stimuli. *J. abnorm. Psychol.* **68**, 221–226.

LUDWIG, A., STILSON, D., WOOD, B. & DOWNS, M. (1963). Further studies in audition in schizophrenia. *Am. J. Psychiat.* **120**, 70–71.

MCGHIE, A. & CHAPMAN, J. (1961). Disorders of attention and perception in early schizophrenia. *Br. J. med. Psychol.* **34**, 103–116.

MCGHIE, A., CHAPMAN, J. & LAWSON, J. (1965). The effect of distraction on schizophrenic performance. I. Perception and immediate memory. *Br. J. Psychiat.* **111**, 383–390.

OSGOOD, C., SUCI, G. & TANNENBAUM, P. (1957). *The Measurement of Meaning*. Urbana, Ill.: University of Illinois Press.

PAYNE, R. W. & FRIEDLANDER, D. (1962). A short battery of simple tests for measuring overinclusive thinking *J. ment. Sci.* **108**, 362–367.

SOMMER, R., DEWAR, R. & OSMOND, H. (1960). Is there a schizophrenic language? *Archs gen. Psychiat.* **3**, 665–673.

SPENCE, J. & LAIR, C. (1964). Associative interference in the verbal learning performance of schizophrenics and normals. *J. abnorm. soc. Psychol.* **68**, 204–209.

THORNDIKE, E. & LORGE, I. (1944). *Teacher's Word Book of 30,000 Words.* New York: Columbia University.

VENABLES, P. (1969). Sensory aspects of psychopathology. In J. Zubin & C. Shagass (eds.), *Neurobiological Aspects of Psychopathology.* New York: Grune & Stratton.

Manuscript received 10 *December* 1969

Revised manuscript received 21 *July* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 253–259
Printed in Great Britain

The Heterogeneity of Personality among Alcoholics

BY KENNETH B. STEIN
University of California, Berkeley

VITALI ROZYNKO AND LAWRENCE A. PUGH
Mendocino State Hospital

The purpose of the investigation was to discover whether multiple personality patterns could be derived from selected personality dimensions among a large sample of alcoholics. Subjects were 302 males hospitalized for a chronic drinking problem. The following six personality dimensions served as the independent variables: socialization, motoric expressiveness, ideational expressiveness, social introversion, bodily concern, and suspicion and mistrust. Utilizing the OTYPE computer program, 11 types or score patterns were derived across the six dimensions. Each of the types was found to be highly homogeneous and capable of differential predictions in relation to a host of psychological and social variables. These significant relationships then served as definitional characteristics of the types. Implications are discussed in relation to aetiology and treatment.

The purpose of the present study is to investigate the heterogeneity of personality among hospitalized chronic alcoholics. Although there are various approaches to the study of the personality of problem drinkers, we have chosen the empirical typological one based upon patterning across personality dimensions.

We selected six variables which are applicable to the study of personality generally and relevant to problem drinking in particular. The first is *socialization*, which has broad conceptual implications and has been discussed in relation to a variety of social and psychological phenomena (cf. Clausen, 1968; Gough, 1960). Although it is expected that chronic alcoholics on the average are more poorly socialized than the population at large, a wide range of individual variation is anticipated which should contribute to the derivation of personality patterns. The second and third dimensions are two independent forms of expressiveness termed *motoric* and *ideational*, which represent preferred modalities and orientations in relating to the world of objects, including the self. Previous work and theoretical discussions relating to these expressive forms can be found in Miller & Swanson (1960) and in the work by Stein and his colleagues (Stein, 1967, 1968*b*; Stein & Craik, 1965; Stein & Lenrow, 1970). Differential preference in expressive style among the alcoholics should have relevance in subsequent studies on modalities of treatment, namely insight and cognitive on the one hand and physical and behaviouristic on the other. *Social introversion*, the fourth variable, has been central to the theory and research of various personality investigators (e.g. Eysenck, 1960). One of the implications of this concept is that it should have relevance to drinking patterns, i.e. solitary or social. The fifth dimension is *bodily concern*, which relates not only to bodily preoccupation but also to emotional well-being. It is our expectation that alcoholics generally would show greater intensity on this dimension than the normal population but at the same time disclose sufficient individual variation to contribute to a meaningful typology. The final dimension is *suspicion and mistrust*, a characterological feature signifying over-alertness and a readiness to project and externalize personal shortcomings. It has significance for the type of relationships formed and the ability for self-scrutiny and behaviour change.

METHOD

Subjects. The subjects consisted of 302 successive admissions to a male alcoholic ward of a California state hospital. These were almost all (90 per cent) voluntary admissions, with a mean age of 45·16 years and a mean education of 11·65 years. Their problem is chronic, the average length of time of uncontrolled drinking being 12·69 years.

Procedures. The MMPI (Hathaway & McKinley, 1951) was administered to the patients and was scored for the standard validity and clinical scales as well as number of multiple marks (mm), questions left blank (BB), Welsh's A and R scales, and Barron's Es. The seven TSC scales (Stein, 1968*a*; Tryon, 1967*b*) were also scored. These latter scales were derived from a cluster analysis of the 550 items of the MMPI and hence constitute pure or non-overlapping item scales. The three most independent of these scales, namely *social introversion* (I), *body complaints* (B) and *suspicion* (S), constitute three of the six personality dimensions of the present study. The remaining TSC scales are *depression* (D), *resentment* (R), *autism* (A) and *tension* (T).

A second administered measure was the *socialization* (So) scale from the California Psychological Inventory (Gough, 1957), a scale which has been widely used in personality research. This scale constituted our measure of the socialization dimension.

The final instrument was the Motoric-Ideational Activity Preference Scale (MIAPS) (Stein & Craik, 1965; Stein, 1967, 1968*b*). This scale supplied scores for the two expressive dimensions, the motoric and the ideational.

In addition to the above personality variables, the following social data were obtained: age, education, marital status, religion, family size, number of marriages, number of children, living arrangements at the time of admission, work history in past year (number of months worked), years of uncontrolled drinking, drinking habits and patterns, whether a member of Alcoholics Anonymous (AA), whether a voluntary or committed admission, and estimated length of hospital stay (weeks). In a post-testing follow-up two years later the number of rehospitalizations were recorded.

The derivation of personality patterns of alcoholics and prediction from these types to a set of non-defining variables involved the use of recently developed computer programs designed for pattern analysis and prediction. The OTYPE program (Tryon, 1967*a*) is an efficient and relatively inexpensive computer method for discovering a finite set of patterns (types) of subjects across dimensional scores.

The prediction program called 4CAST (Tryon, 1967*b*) operates on the Monte Carlo principle. For each type on each variable to be predicted, 300 successive random samples, each equal in number to the type, are selected from the total sample and the distribution of these means is compared with that of the type. In this fashion a direct probability is presented by the calculation of the proportion of random sample means that fall above and below the mean of the type. This method of prediction avoids interpolated ratios, requires no statistical assumptions, and is a highly stringent test of significance.

RESULTS

Characteristics of the total sample. The greater percentage of subjects were divorced (38·1 per cent), 25·5 per cent were married and 15·2 per cent were single. An additional 15·9 per cent were separated from their wives and 5·3 per cent were widowers. The majority of subjects professed to be Protestants (54·9 per cent), 38·7 per cent were Catholics, fewer than 1 per cent were Jews, and 2·7 per cent had

Table 1. *Means, sigmas, and intercorrelations of the six personality dimensions among 302 hospitalized alcoholic patients*

			Correlations					
Variable	Mean	S.D.	TSC I	TSC B	TSC S	CPI So	Motor	Ideate
TSC Introversion (I)	11·12	6·40						
TSC Body Complaint (B)	9·49	6·61	0·42					
TSC Suspicion (S)	14·41	5·65	0·35	0·43				
CPI Socialization (So)	28·21	6·35	−0·26	−0·31	−0·36			
Motoric (Motor)	60·98	15·39	−0·12	−0·16	0·06	0·01		
Ideational (Ideate)	50·42	16·07	−0·06	−0·06	−0·15	0·06	0·38	

no religious affiliation. There were 40·1 per cent who were assessed as lone drinkers, 42·8 per cent as social drinkers, and only 17·2 per cent as both lone and social drinkers. The largest percentage (48·8 per cent) disclosed a pattern of 'binge drinking' and 42·8 per cent were steady drinkers. The remaining 8·4 per cent showed a mixed pattern of both binge and steady drinking. The mean age at which the patients were first married was 23·5 years. At the time of hospital admission the majority (52·8 per cent) were living alone, 22·3 per cent with mate, 11·7 per cent with relatives, and 13·2 per cent with friends.

Although the primary interest is in hospitalized alcoholic patients and their classification by personality patterns, some comparisons on the six dimensions between the present study sample and normative data are of interest. Table 1 presents the means, standard deviations and intercorrelations of the six personality dimensions. The TSC norms can be found in a recent published study (Stein, 1968*a*, Table 4, p. 98). The mean of each of the three dimensions of the normative sample is significantly lower, indicating greater social introversion, more bodily complaints, and increased suspicion on the part of the alcoholic study sample. The So scale mean score is low and is more like that of less socialized, delinquent or imprisoned males (Gough, 1957, Table 3, p. 34) than normals or even hospitalized psychiatric patients. When compared with a normal sample, the alcoholics do not differ on the ideational expressive dimension but do show greater intensity on the motoric scale (Stein, 1964).

As expected, there are a number of significant correlations between the dimensions but these tend to be modest and account for small proportions of the variance. The magnitude of intercorrelations of the TSC scales is somewhat greater than in the original study (cf. Stein, 1968*a*, Table 1, p. 90). The So scale is negatively related to each of the TSC scales, which is hardly surprising, but is not related to either motoric or ideational expressiveness. The modest covariation of the two expressive dimensions indicates that they are not opposite ends of a single dimension. Generally, none of the correlations among the six dimensions is of a size which would interfere with or restrict the derivation of types to any great extent.

Alcoholic type derivation. Eleven types emerged from the OTYPE analysis. Table 2 presents the mean dimensional scores for each type in standard scores in which the mean is 50 and the standard deviation is 10. The goodness of fit can be judged from several statistics in the table. First, an inspection of the standard deviations indicates that all of them are below 10, the standard deviation of the entire sample. The eta summarizes the homogeneity of scores within each dimension across the types. These scores range from 0·77 to 0·83, indicating that the variation of scores within types for each dimension is quite narrow. A second measure of goodness of fit is the *H* or overall partial eta across dimensions within each type, that is, the mean of the six partial etas for each type. These *H*s range from 0·72 to 0·82, again reflecting a high degree of pattern similarity for subjects within types. Also, it should be noted that 288 of the 302 subjects were accounted for with the 11 types (95·4 per cent). The 14 unassigned subjects are individuals with unique patterns who do not have a sufficiently close fit with any of the types. Less than 5 per cent is a small loss in order to achieve satisfactory homogeneity.

Table 2. Standard score patterns of alcoholic types

Alcoholic type	*TSC Introv.*	*TSC Body*	*TSC Susp.*	*CPI So*	*Motoric*	*Ideational*	*H*	*n*
1. *Socialized and emotionally undisturbed*								
Mean	40·1	42·5	37·1	57·4	45·1	45·3	0·82	37
S.D.	4·6	4·7	5·3	6·3	7·2	6·1		
2. Poorly socialized								
Mean	49·4	48·4	47·4	37·9	43·2	47·3	0·80	16
S.D.	7·3	6·1	5·6	5·1	4·8	7·1		
3. Introverted								
Mean	58·4	48·1	41·9	54·9	43·5	47·9	0·80	21
S.D.	4·8	5·8	4·6	6·9	5·8	7·5		
4. Socialized but mistrusting								
Mean	45·6	46·8	53·5	58·6	47·2	44·3	0·82	36
S.D.	4·5	5·5	5·6	5·7	7·3	5·2		
5. Ideational								
Mean	40·6	41·8	40·2	47·5	50·1	62·5	0·81	23
S.D.	4·7	3·4	6·8	5·8	7·2	6·1		
6. Poorly socialized motoric								
Mean	47·5	50·5	52·1	41·9	65·5	51·8	0·76	23
S.D.	5·5	6·9	8·3	6·5	4·1	6·9		
7. Socialized high expressive								
Mean	42·7	41·8	44·4	62·5	62·2	59·9	0·78	24
S.D.	6·9	4·3	6·9	5·4	6·4	7·1		
8. Suspicious and poorly socialized								
Mean	47·7	53·1	60·3	40·9	49·2	45·6	0·82	29
S.D.	6·5	6·7	3·8	5·8	5·2	5·7		
9. Emotionally disturbed low expressive								
Mean	55·9	66·8	57·2	47·3	40·7	42·7	0·78	28
S.D.	7·3	6·0	4·6	7·5	5·6	5·8		
10. Emotionally disturbed ideational								
Mean	60·8	62·3	57·8	46·1	53·9	64·1	0·72	25
S.D.	6·6	6·7	6·8	6·8	6·6	8·0		
11. Suspicious and introverted								
Mean	63·5	46·7	56·2	48·3	49·9	43·9	0·81	26
S.D.	5·1	4·1	5·8	7·0	6·8	6·1		
Eta	0·82	0·83	0·81	0·78	0·78	0·77		
n								288

The appendix contains a table of the means and proportions of variables for each type and a designation of which variables are significantly divergent from the sample as a whole.* The significant variables serve as additional definers for type description. The 55 variables in this table include the MMPI and TSC scales, social and background factors, and the number of post-testing rehospitalizations. The tentative names assigned to the types are based upon the score pattern, even though more 'catchy' or descriptive titles may eventually emerge from the validity data. Space does not allow for a narrative description of each type, so these also have been included in the appendix.

DISCUSSION

The evidence of heterogeneity of personality factors among alcoholics is quite striking. This sample includes typological subgroups that range from introverted, withdrawn and shy persons to those who are outgoing and interpersonally assertive. There are those who are plagued by bodily concerns and emotional disturbance and those who endorse very few such complaints. Suspicion, level of socialization and motoric and ideational expressiveness also show a wide range and diversity among the types. The range of scores of types for each personality dimension is minimally two standard deviations. When it is considered further that each type is relatively homogeneous in its score pattern, such a wide range of scores on each personality dimension is a compelling refutation of a conceptualization of a single, homogeneous 'alcoholic personality'. Although this is not a new contention (e.g. Armstrong, 1958; Nurse, 1965; Mogar *et al.*, 1970), much research and clinical investigation still appear to be guided by this notion of homogeneity.

There is a value in keeping the heterogeneity conception in mind in that it provides an orientation to the research which is done. There are at least two directions that have been of greatest interest to investigators. The first of these is the area of aetiology. The assumption of heterogeneity calls not only for a search for multiple causal factors but also for a consideration of differential causal patterns. The second prominent direction of investigation has been in the area of treatment or behaviour change. An orientation of heterogeneity shifts the question from 'What kind of treatment for the alcoholic personality?' to 'What different kinds of treatment for different types of problem drinkers?' Such a reformulated question is more conducive to a systematic investigation of the interaction of types and treatment modes.

The directions for future research relate to the two areas mentioned above as well as to a third. First, these types would serve as the frames of reference from which an investigation of the multiple aetiological factors would follow. In order to avoid confounding causes and consequences of drinking problems, longitudinal as well as in-depth adult studies need to be done, as, for example, in the work of Jones (1968) and McCord & McCord (1960). It would also be important in this area to observe whether the same personality types can be located among non-problem drinkers and then to examine the similarities and differences.

---

* An appendix containing the table of means and proportions of variables for each of the 11 types as well as a brief narrative description of each type is available from the first author at no charge.

The second direction is in treatment or behaviour change. These 11 types would serve as the points of reference in the study of relations of responsiveness to differential modes of treatment. The growing alcohol problem makes it as important to discover adequate modes of treatment for application to those for whom it will be most effective as it is to discover aetiological factors for purposes of prevention.

The third direction is in systematic exploration of other personality dimensions in order to discover convergence in sets of personality patterns produced. Such additional dimensions selected for study will probably prove useful if they are based upon theoretical formulations meaningfully related to the problem of drinking (cf. Sanford, 1968).

REFERENCES

Armstrong, J. D. (1958). Search for the alcoholic personality in understanding alcoholism. *Ann. Am. Acad. pol. soc. Sci.* **135**, 40–47.

Clausen, J. A. (ed.) (1968). *Socialization and Society*. Boston: Little, Brown.

Eysenck, H. (1960). *The Structure of Human Personality*. New York: Macmillan.

Gough, H. G. (1957). *Manual for the California Psychological Inventory*. Palo Alto, Calif.: Consulting Psychologists Press.

Gough, H. G. (1960). Theory and measurement of socialization. *J. consult. Psychol.* **24**, 23–30.

Hathaway, S. R. & McKinley, J. C. (1951). *The Minnesota Multiphasic Personality Inventory Manual*, rev. ed. New York: Psychological Corporation.

Jones, M. C. (1968). Personality correlates and antecedents of drinking patterns in adult males. *J. consult. clin. Psychol.* **32**, 2–12.

McCord, W. & McCord, J. (1960). *Origins of Alcoholism*. Stanford, Calif.: Stanford University Press.

Miller, D. R. & Swanson, G. E. (1960). *Inner Conflict and Defence*. New York: Holt.

Mogar, R. E., Wilson, W. M. & Helm, S. T. (1970). Personality subtypes of male and female alcoholic patients. *Int. J. Addict.* **5**, 99–113.

Nurse, A. R. (1965). Personality changes relative to therapy of alcoholics. (A final report on NIMH Grant 6360. Bethesda, Maryland: National Institute of Mental Health.)

Sanford, N. (1968). Personality and patterns of alcoholic consumption. *J. consult. clin. Psychol.* **32**, 13–17.

Stein, K. B. (1964). Some correlates of the motoric and ideational sub-clusters among college students. (Unpublished report; University of California, Berkeley.)

Stein, K. B. (1967). Correlates of the ideational preference dimension among prison inmates. *Psychol. Rep.* **21**, 553–562.

Stein, K. B. (1968*a*). The TSC scales: the outcome of a cluster analysis of the 550 MMPI items. In P. McReynold (ed.), *Advances in Psychological Assessment*. Palo Alto, Calif.: Science & Behavior Press.

Stein, K. B. (1968*b*). Motoric and ideational dimensions: a study of personality styles among prison inmates. *Psychol. Rep.* **22**, 119–129.

Stein, K. B. & Craik, K. H. (1965). Relationship between motoric and ideational activity preference and time perspective in neurotics and schizophrenics. *J. consult. Psychol.* **29**, 460–467.

Stein, K. B. & Lenrow, P. (1970). Expressive styles and their measurement. *J. Person. soc. Psychol.* **16**, 656–664.

Tryon, R. C. (1967*a*). Person clusters on intellectual abilities and MMPI attributes. *Multivar. behav. Res.* **2**, 5–34.

Tryon, R. C. (1967*b*). Predicting individual differences by cluster analysis: Holzinger abilities and MMPI personality attributes. *Multivar. behav. Res.* **2**, 325–348.

Manuscript received 2 *March* 1970

Revised manuscript received 14 *September* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 260–266
Printed in Great Britain

A Study of the Reliability Implications of Lie Scores in the Junior Eysenck Personality Inventory

By WILLIAM ORWIN

Department of Social Psychology, University of Sussex

Two groups of 8-year-old children were each administered the JEPI on two occasions, using the headmaster of their school and the present author as test administrators. When tested by the headmaster the boys recorded significantly higher Neuroticism scores and significantly lower Lie scores; the girls' scores did not differ between testers. Three different tests of the reliability and consistency of the boys' N scores each failed to support predictions arising from the assumption that low L scores are associated with more reliable or more consistent N scores.

The Lie (L) scale of the JEPI (and of the EPI) was constructed and included in the inventory as a means of 'screening out' invalid protocols. Attempts to test empirically that high L scores are associated with invalid Neuroticism (N) and Extraversion (E) scores have been methodologically unsound. The numerous 'fake good' and 'warning' studies (e.g. Braun & Gomez, 1966; Eysenck & Eysenck, 1963; Power, 1968; Power & O'Donovan, 1969; Star, 1962) have, by virtue of their between-groups experimental designs, yielded information strictly on *treatment effects*. The contradictory conclusions drawn from such studies have arisen from attempts to interpret group treatment effects in terms of estimates of the validity of individual scores. The validity *implications* of L scores are in any case difficult to discern, since very little is known about the validity of the JEPI scales themselves (S. B. G. Eysenck, 1965, p. 13), and since there is no information as to the factor loadings of L-scale items on the E and N scales (S. B. G. Eysenck, personal communication, 1969).

The hypothesis that L scores have *reliability* implications for N scores can, however, be tested. Thus Eysenck *et al.* (1966) claim to have shown that, under 'fake good' instructions, N scores tend to be less reliable than scores obtained normally, and suggest further that this decrement in reliability is 'in good accord with the changes in means' of N and L scores produced by the 'fake good' instructions under which L scores were inflated and N scores depressed (pp. 89–90). The observation of a 'good accord' derives, of course, from the two-fold assumption that the L scale measures a 'faking good' tendency and, as a corollary, that 'fake good' instructions will exacerbate whatever degree of invalidity and unreliability may already be present in normally obtained N scores. It should be noted, however, that the test–retest correlations in this study were not proper estimates of reliability since scores on occasion 1 were taken from a 'no treatment' condition, while scores on occasion 2 were taken from a treatment condition. The present study re-examines the hypothesis that L scores have reliability implications for N scores, and then extends the hypothesis to include the notion of 'consistency', which is argued to be an elaboration of reliability of the sort that has a greater relevance for estimates of validity (Dempsey, 1965).

METHOD

The JEPI was administered to two classes of 8-year-old children in a small, city primary school. Both classes were given the JEPI on two occasions, these being exactly 14 days apart. The cooperation of the headmaster of the school was secured on the hypothesis that L scores obtained by him would differ from those obtained by the writer, who, of course, was a stranger to the children. The headmaster was fully briefed in the normal administration procedure, and was strongly urged to avoid departing from it. All items were read aloud to the classes, and the headmaster and the present author agreed, before each of them administered the test, which 'difficult' items should be elaborated and in what terms. Since the two groups of children were 'parallel classes' which had been formed by random assignment of children to them, a cross-over design (Cochran & Cox, 1957; Maxwell, 1958) was set up. Thus, on day 1, the test was administered to class 1 by the headmaster (H) and to class 2 by the present author (A). On day 2, two weeks later at exactly the same time of day, A administered the test to class 1, while H administered the test to class 2. This design permits an analysis of within-subjects shifts in JEPI scores, where each subject serves as his own control (cf. the control *groups* in the 'fake good' studies), allows a 'between occasions' source of variation to be partitioned, and satisfactory measures of reliability to be calculated.

RESULTS AND DISCUSSION

Tables 1 and 2 show the means and standard deviations of the E, N and L scores for the two groups, for boys and girls respectively. The data are arranged in accordance with the cross-over design. Tables 3 and 4 show analyses of variance of the N and L scores respectively for the boys. A log transformation was applied to the data since the standard deviations tended to be proportional to the means.

The results show that the mean N score was significantly higher when H tested the boys than when A tested them, and that the mean L score was significantly lower when H tested the boys. Variation due to occasions of testing and to groups was not significant for either set of scores. By similar analysis of variance, there emerged no main effect due to testers in the case of the boys' E scores ($F=0{\cdot}62$; d.f. $=1, 36$, n.s.). Nor were there any significant main effects in the girls' scores.

Table 5 shows a comparison between the E, N and L scores of the present sample of boys with those of the group of boys of the same age used by S. B. G. Eysenck

Table 1. *Means and standard deviations of boys' E, N and L scores*

	Day 1			Day 2		
	E	N	L	E	N	L
Group 1 ($n=19$)		H			A	
Mean	17·47	14·84	4·37	17·16	11·74	6·26
S.D.	3·10	3·22	2·11	3·81	5·90	2·45
Group 2 ($n=19$)		A			H	
Mean	15·42	12·37	6·68	16·32	14·84	5·74
S.D.	3·85	4·46	2·42	4·64	5·07	3·09

Table 2. *Means and standard deviations of girls' E, N and L scores*

	Day 1			Day 2		
	E	N	L	E	N	L
Group 1 ($n=11$)		H			A	
Mean	16·09	16·55	6·73	15·55	13·82	7·36
S.D.	3·42	4·11	1·79	3·75	4·89	3·07
Group 2 ($n=12$)		A			H	
Mean	15·89	13·89	6·78	14·67	12·00	6·67
S.D.	4·31	5·04	2·97	4·09	5·29	3·61

Table 3. *Analysis of variance of boys' N scores*

Source of variation	S.S.	D.F.	M.S.	*F*	*P*
Between treatment groups					
Between days	0·0316	1	0·0316	1·24	n.s.
Between testers	0·2969	1	0·2969	11·69	< 0·002
Between groups	0·0114	1	0·0114	0·45	n.s.
Total	0·3399	3			
Within treatment groups					
Between subjects within treatment groups	1·9323	36	0·0537	2·11	< 0·01
Residual (error)	0·9148	36	0·0254		
Grand total	3·1869	75			

Table 4. *Analysis of variance of boys' L scores*

Source of variation	S.S.	D.F.	M.S.	*F*	*P*
Between treatment groups					
Between days	0·0127	1	0·0127	0·60	n.s.
Between testers	0·3725	1	0·3725	17·74	< 0·001
Between groups	0·0708	1	0·0708	3·37	n.s.
Total	0·4560	3			
Within treatment groups					
Between subjects within treatment groups	3·0658	36	0·0852	4·06	< 0·001
Residual (error)	0·7577	36	0·0210		
Grand total	4·2795	75			

(1965) as her standardization sample. For the purpose of these comparisons, since there were no significant differences between groups, the scores of the two groups were combined. Table 5 shows that, when tested by H, the boys gave significantly higher N scores and significantly lower L scores than were obtained in the standardization samples, and that the E scores did not differ. When tested by A, however, the same sample of boys gave mean E, N and L scores which were extremely close to the standardization means.

Table 5. *Comparisons between sample means and standardization norms*

		Present sample			Standardization sample				
		Mean	S.D.	*n*	Mean	S.D.	*n*	*t*	*P*
H	E	16·89	3·940	38	16·67	3·128	433	0·33	n.s.
	N	14·84	4·188	38	11·52	4·848	433	4·62	<0·001
	L	5·05	2·701	38	6·67	2·662	138	3·28	<0·001
A	E	16·29	3·876	38	16·67	3·128	433	0·59	n.s.
	N	12·05	5·167	38	11·52	4·848	433	0·61	n.s.
	L	6·47	2·402	38	6·67	2·662	138	0·44	n.s.

What thus emerges is an experimentally induced depression in L scores and an experimentally induced elevation of N scores in a sample of 38 boys. Further, there emerges a fully adequate control *condition* within the same subjects against which to compare the properties of the abnormal scores.

Split-half reliability. On the assumption that the L scale (whatever it is measuring) has reliability implications for N scores, and given the significantly lower L scores obtained under the H condition of testing, one prediction that can be made is that split-half reliabilities of N scores obtained in the H condition should be higher than those of the N scores obtained in the A condition. The data yield the following split-half reliability coefficients for the N scores (incorporating the Spearman–Brown correction for reduced length of test): when tested by H, 0·704; when tested by A, 0·821. The raw correlation coefficients do not differ significantly: $r=0{\cdot}543$ and 0·697 respectively; $t=1{\cdot}03$. The prediction that low L scores are associated with more reliable N scores fails to receive support by this test. Indeed, the order of magnitude of the coefficients is in the direction opposite to that required by the prediction.

An elaboration of this procedure is to compare the split-half reliability coefficient of N scores associated with low L scores with that for N scores associated with high L scores, irrespective of administration condition. Again, on the assumption that the L scale has reliability implications for N scores, the prediction arises that these split-half reliability coefficients should be high and low respectively. The cut-off points taken to separate obtained high and low L scores were scores of 9 and 4 respectively. Taken irrespective of administration condition, there is a total of 76 L scores. Table 6 shows that the mean low L score of the present sample is separated from the mean high L score by more than two standard deviations (one either side of the mean) of the standardization distribution.

Table 6. *Split-half reliabilities of N scores associated with high and low L scores*

	L scores			N scores			
	Mean	S.D.	*n*	Mean	S.D.	Correlation	Reliability
Low L scores	2·93	1·052	28	15·25	2·469	+0·645	+0·784
High L scores	9·46	0·519	13	11·15	2·611	+0·747	+0·855

The table also shows that the correlations between the two sets of N scores do not differ significantly, and that the split-half reliability coefficients are not only acceptably high but also, again, show an order of magnitude which is opposite to the one predicted. By this analysis, then, the prediction that low L scores are associated with more reliable N scores fails to receive confirmation.

Consistency of single scores. The split-half reliability coefficient is, of course, a fairly crude measure of consistency since it utilizes only gross frequency counts. It could be argued that L scores detect not simply discrepancies between endorsement frequencies on the N scale but individual inconsistencies *within identical* endorsement frequencies. In other words, the trans-scalar reliability (or validity) assumption of the presence of the L scale might predict that low L scores are associated with N-scale endorsement patterns which indicate more consistent *syndromes* of neuroticism, however high or low the total N scores may be.

A statistic devised by Dempsey (1965) may be used in connexion with this prediction. Dempsey offers an index of the 'difficulty' level of a personality item which takes the form: $d_k=(\sum s_{jk})/n$, where d_k denotes the difficulty of item k, s_{jk} is the scale score of any subject who endorses (passes) item k, and n is the number of subjects so passing. Dempsey elaborates the use of this statistic to detect 'invalid protocols': 'When the difficulty levels of the items on a personality scale are known, inconsistent performances can be readily identified. In particular, subjects who show correct responses to difficult items but not to easy items may be suspected of faking, of being confused as to instructions, of not belonging to the population for whom the scale was devised, or of some other invalidating factor' (p. 146). A clear prediction emerges: given the assumption that high L scores are associated with inconsistent N scores, and given the significantly lower L scores obtained in the H condition, there should be a higher mean N-scale consistency score in the H condition than in the A condition. Accordingly, a measure of N-scale consistency was computed for each subject under each testing condition. The four sets of scores were transformed into logs since the standard deviations tended to be proportional to the means. A cross-over analysis of variance of consistency scores was performed and is given in Table 7.

Table 7. *Analysis of variance of boys' N-scale S_{dk} scores*

Source of variation	S.S.	D.F.	M.S.	*F*	*P*
Between treatment groups					
Between days	0·0031	1	0·0031	0·19	n.s.
Between testers	0·0475	1	0·0475	2·84	n.s.
Between groups	0·0001	1	0·0001	0·01	n.s.
Total	0·0507	3			
Within treatment groups					
Between subjects within treatment groups	1·3171	36	0·0366	2·19	$<0·01$
Residual (error)	0·6026	36	0·0167		
Grand total	1·9197	75			

The analysis shows that the N scores of individual subjects were no more consistent in the H condition than in the A condition. By this test, therefore, the prediction that low L scores are associated with more consistent N scores fails to receive support.

Consistency of shift. A final measure of individual consistency is one which assesses the consistency a subject shows in his *shift* to greater neuroticism in the H condition. Thus, regardless of gross amount of increase in neuroticism between testers, if subjects' N scores are reliable, one would expect the subjects to endorse a substantial number of common N items between testers, such that their higher N scores comprise (*a*) items representing their 'normal' N scores, plus (*b*) some items extra to these but just sufficient in number to make up their higher N scores. A statistic which summarizes this set-theoretical information, and which makes no assumption as to unidimensionality of the obtained N scores, is given by: $C = \sum i_c/[\sum i_c + \sum i_s]$, where $\sum i_c$ denotes the number of common items endorsed between conditions, and $\sum i_s$ denotes the number of endorsed items specific to the condition in which the subject gives the smaller total score.

This C statistic was calculated for all 38 subjects, and assumed a mean value of 0·77 with a standard deviation of 0·163. Given the assumption that low L scores are associated with greater consistency of between-condition shifts in N scores (which would indicate a higher stability of recorded neuroticism syndromes), a prediction emerges that there should be a significant negative correlation between C scores and raw L scores, regardless of condition of testing. Obtained correlations were as follows: for the H condition, $r = +0{\cdot}005$, and, for the A condition, $r = +0{\cdot}002$. By this measure of consistency, therefore, the prediction that low L scores are associated with more consistent N scores fails to receive support.

In summary, this study suggests that the L scale has no reliability or consistency implications for the N scale: the reliability and consistency of the present N scores is just as high (or low) when associated with high L scores as when associated with low L scores.

The present study was carried out during the tenure of a pre-doctoral scholarship for training in research methods from the Medical Research Council, held while the author was at the Institute of Experimental Psychology, University of Oxford. The author is grateful to Mr B. Durbin, Headmaster of Cowley, St Christopher Junior School, Oxford, for his generous cooperation; and to Mr A. D. Chalmers and Mr I. Vilkkumaa for their helpful comments on an early draft of this paper.

REFERENCES

Braun, J. R. & Gomez, B. J. (1966). Effects of faking instructions on the Eysenck Personality Inventory. *Psychol. Rep.* **19**, 388–390.

Cochran, W. G. & Cox, G. M. (1957). *Experimental Design*, 2nd ed. New York: Wiley.

Dempsey, P. (1965). The concept of item difficulty in personality measurement: a note on Hanley's formulation. *Educ. psychol. Measur.* **25**, 143–147.

Eysenck, S. B. G. (1965). *Manual of the Junior Eysenck Personality Inventory*. London: University of London Press.

Eysenck, S. B. G. & Eysenck, H. J. (1963). An experimental investigation of 'desirability' response set in a personality questionnaire. *Life Sci.* **2**, 343–355.

EYSENCK, S. B. G., SYED, I. A. & EYSENCK, H. J. (1966). Desirability response set in children. *Br. J. educ. Psychol.* **36**, 87–90.

MAXWELL, A. E. (1958). *Experimental Design in Psychology and the Medical Sciences.* London: Methuen.

POWER, R. P. (1968). Simulation of stable and neurotic personalities by subjects warned of the presence of lie scales in inventories. *Br. J. Psychol.* **59**, 105–109.

POWER, R. P. & O'DONOVAN, D. (1969). Detection of simulation on the MPI by subjects given the rationale of the Lie scale. *Br. J. Psychol.* **60**, 535–541.

STAR, K. H. (1962). Ideal-self response set and Maudsley Personality Inventory scores. *Psychol. Rep.* **11**, 708.

Manuscript received 24 *March* 1970

Revised manuscript received 29 *June* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 267–270
Printed in Great Britain

'Psychological' Constructs, Thought-Process Disorder and Flattening of Affect

BY F. M. McPHERSON, FELICITY BUCKLEY AND JOAN DRAFFAN

University Department of Psychiatry, Royal Edinburgh Hospital, Edinburgh

The frequency with which schizophrenics *spontaneously* use 'psychological' constructs, as opposed to other categories of construct, when describing photographs of people is significantly and negatively correlated with the extent of disorder ('looseness' and 'inconsistency') shown when they are *required* to use 'psychological' constructs, on the Bannister–Fransella test.

This study relates disorders of the *structure* and of the *content* of the conceptual systems of schizophrenic patients.

Schizophrenics with clinical signs of thought-process disorder are characterized by disturbances of their conceptual structure (Bannister, 1960, 1962). On a version of the Repertory Grid test (Kelly, 1955; Bannister & Mair, 1968), Bannister & Fransella (1966) showed that the intercorrelations between constructs (concepts) are abnormally low ('looseness') and that the pattern of intercorrelations is unstable ('inconsistency'). Thought-process disorder does not appear to affect all constructs equally. 'Psychological' constructs—those referring to the realm of human thought, feeling, personality, abilities, etc.—are much more disordered than, for example, constructs descriptive of the physical properties of people and objects (Bannister & Salmon, 1966; Salmon *et al.*, 1967; McPherson & Buckley, 1970).

The measures used in these studies required the subject to describe elements (people, objects) according to constructs provided by the tester. More usually in everyday situations, however, a person is free to construe events in terms of any constructs from among those in his repertoire. Dixon (1968) and McPherson *et al.* (1970*a*, *b*) assessed this 'spontaneous' use of constructs by asking subjects to describe differences between people in photographs. All three studies demonstrated another abnormality associated with 'psychological' constructs. Schizophrenics with the clinical signs of 'flattening of affect' made much less frequent use of 'psychological' constructs than did schizophrenics without affective flattening, and normals; they did not differ in their use of several other categories of construct, such as those referring to the activities or physical features of the people.

Taken together, these two sets of results suggest that the extent to which a schizophrenic patient, when construing other people, emphasizes their 'psychological' characteristics, as opposed to their physical features etc., is influenced by his ability to use 'psychological' constructs in an orderly manner, i.e. the more disordered these constructs are, as measured by the Bannister–Fransella test, the less they will be used 'spontaneously', as assessed by Dixon's procedure.

METHOD

Subjects. Forty-six schizophrenic patients were tested; these comprised all the testable patients in the admission wards of a psychiatric hospital, on two occasions, who were aged 67 or under; scored 15 or more on Mill Hill Synonym Selections (Senior); had an undisputed diagnosis of schizophrenia; and were exhibiting active psychotic signs and symptoms. Twenty-five patients (seven females, 18 males) were acute, i.e. within two years of first admission to hospital; they had a mean age of 26·4 years (S.D. 10·8). The 21 chronic patients (nine females, 12 males) had a mean age of 38·8 years (S.D. 13·2). The mean Synonym Selection test score of the acutes was 28·3 (S.D. 5·3) and of the chronics 26·0 (S.D. 4·1).

Measures. The Bannister–Fransella test (1966) was administered in the usual way. The subjects' spontaneous use of constructs was assessed in the way described by Dixon (1968) and McPherson *et al.* (1970*a*, *b*). The subject was asked to describe all the differences he could observe between the *people* in a pair of photographs, of people engaged in various activities; 3 min. were given for this. Five pairs were presented (shown in Dixon, 1968), the instructions being repeated before each pair. The responses were tape-recorded and content-analysed by two scorers, who worked independently and had no knowledge of the Bannister–Fransella test scores. For every patient, the scorers decided for each pair of photographs whether or not each of five categories of construct had been used. These five were:

Activities: constructs which distinguished between the people according to their activities or stance, e.g. 'reading a book – not reading', 'standing – sitting'.

Psychological: those which referred to the personality, emotional state or abilities of the people, e.g. 'looks honest – dishonest', 'happy – sad', 'intelligent – unintelligent'.

Physical: those which described the physical characteristics of the people, including their age, e.g. 'well-built – frail', 'old – young'.

Social: those which referred to their social class, occupation or nationality, e.g. 'a working man – a professional type', 'looks foreign – looks British'.

Non-human: some constructs referred to the clothes the people were wearing or to inanimate, 'background' objects (although the instructions emphasized differences between the *people*), e.g. 'wearing a scarf – not', 'there is a book in this photo – not in this one'.

For each pair of photographs, each category was given a score of 1 if it was used, and 0 if not. Thus each patient had a score for each category ranging from 0–5, and for all five categories together of 0–25. Each of the five single category scores was expressed as a percentage of the total score to indicate the *relative* use made of each category of construct by the patient.

Inter-scorer reliability was very high. Over the 46 patients, 1150 scoring decisions were made; the scorers disagreed on only seven of these: the mean of the two scores was used in the subsequent analysis.

RESULTS

The acute and chronic groups were analysed separately. The rank order correlation coefficients (Kendall's tau) calculated between the intensity and consistency scores of the Bannister–Fransella test and the percentage frequency of use of the five categories of construct are shown in Table 1.

All four correlations between Bannister–Fransella test scores and the use made of 'psychological' constructs were positive and statistically significant, i.e. the more disordered patients made less use of constructs of this type. None of the other correlations was significant. The only consistent pattern was the four negative correlations with the non-human category, i.e. the more disordered patients tended to make more references to the clothes of the people in the photographs, and (contrary to the instructions) to background objects.

There were no significant correlations between vocabulary level and any of the above scores.

Table 1. *Rank-order correlation coefficients (Kendall's tau) between intensity and consistency scores and use of five categories of construct in acute and chronic schizophrenics*

	Intensity	*P*	Consistency	*P*
Acute ($n=25$)				
Activities	−0·12	n.s.	−0·20	n.s.
Psychological	+0·53	<0·0001*	+0·37	<0·005*
Physical	+0·03	n.s.	−0·14	n.s.
Social	−0·13	n.s.	+0·04	n.s.
Non-human	−0·27	<0·06†	−0·18	n.s.
Chronic ($n=21$)				
Activities	+0·10	n.s.	−0·12	n.s.
Psychological	+0·38	<0·01*	+0·27	<0·05*
Physical	+0·06	n.s.	−0·02	n.s.
Social	−0·19	n.s.	−0·09	n.s.
Non-human	−0·10	n.s.	−0·11	n.s.

* One-tailed test. † Two-tailed test.

DISCUSSION

Those schizophrenic patients who are most disordered when required to use 'psychological' constructs are thus least likely to use them when free to use constructs of any type. A possible explanation is the suggestion by Bannister (1963, 1965), supported by some research evidence, that 'looseness' and 'inconsistency' are responses to repeated invalidation of constructs. According to this view, constructs are hypotheses whose purpose is the prediction and anticipation of events. In the face of continual failure to predict events accurately, the schizophrenic loosens the relationships among his constructs, so that they generate less precise predictions and are consequently less liable to subsequent invalidation. The constructs which a person uses frequently are presumably those which enable him to make accurate predictions, and, conversely, those whose predictive value is low will tend not to be used. Thus Poch (1952) showed that predictive constructs which were experimentally invalidated tended later to be abandoned in favour of other constructs. The thought-disordered schizophrenic, whose 'psychological' construct subsystem is loosely organized and does not permit accurate predictions to be made, may therefore attempt to avoid the consequences of this by ignoring, whenever possible, the personality traits and emotional state of other people, and emphasizing instead 'non-psychological' characteristics about which valid predictions are possible.

In general, the present results, together with those of the studies noted above, suggest that two apparently distinct clinical signs of schizophrenia—thought-process disorder and 'flattening of affect'—may both be associated with a more basic abnormality—the absence of a coherent system for construing psychological states.

We are grateful for the cooperation of the staff of the Royal Edinburgh Hospital. F.M.M. was supported by the Mental Health Research Fund, F.B. by the Social Science Research Council and J.D. by the Medical Research Council. The assistance of all three bodies is gladly acknowledged.

REFERENCES

Bannister, D. (1960). Conceptual structure in thought-disordered schizophrenics. *J. ment. Sci.* **106**, 1230–1249.

Bannister, D. (1962). Personal construct theory: a summary and experimental paradigm. *Acta psychol.* **20**, 104–120.

Bannister, D. (1963). The genesis of schizophrenic thought disorder: a serial invalidation hypothesis. *Br. J. Psychiat.* **109**, 680–686.

Bannister, D. (1965). The genesis of schizophrenic thought disorder: retest of the serial invalidation hypothesis. *Br. J. Psychiat.* **111**, 377–382.

Bannister, D. & Fransella, F. (1966). Grid test of schizophrenic thought disorder. *Br. J. soc. clin. Psychol.* **5**, 95–102.

Bannister, D. & Mair, J. M. M. (1968). *The Evaluation of Personal Constructs*. London and New York: Academic Press.

Bannister, D. & Salmon, P. (1966). Schizophrenic thought disorder: specific or diffuse? *Br. J. med. Psychol.* **39**, 215–219.

Dixon, P. M. (1968). Reduced emotional responsiveness in schizophrenia. (Unpublished Ph.D. thesis, University of London.)

Kelly, G. A. (1955). *The Psychology of Personal Constructs*, vols. 1–2. New York: Norton.

McPherson, F. M., Barden, V. & Buckley, F. (1970*a*). The use of 'psychological' constructs by affectively flattened schizophrenics. *Br. J. med. Psychol.* **43**, 291–293.

McPherson, F. M., Barden, V., Hay, A. J., Johnstone, D. W. & Kushner, A. W. (1970*b*). Flattening of affect and personal constructs. *Br. J. Psychiat.* **116**, 39–43.

McPherson, F. M. & Buckley, F. (1970). Thought-process disorder and personal construct subsystems. *Br. J. soc. clin. Psychol.* **9**, 380–381.

Poch, S. M. (1952). A study of changes in personal constructs as related to interpersonal prediction and its outcome. (Unpublished Ph.D. thesis, Ohio State University.)

Salmon, P., Bramley, J. & Presly, A. S. (1967). The Word-in-Context test as a measure of conceptualization in schizophrenics with and without thought disorder. *Br. J. med. Psychol.* **40**, 253–259.

Manuscript received 26 March 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 271–279
Printed in Great Britain

A Study of Worrying

By SHLOMO BREZNITZ

The Hebrew University, Jerusalem, Israel

Worrying has never been properly studied. It is difficult to induce and to measure and has never been clearly defined. Some view it as neurotic behaviour synonymous with anxiety, whereas others indicate the positive aspects of the 'work of worrying'. Seen mainly as a cognitive process, worrying is defined as a process of gradual involvement with external threat due to its relevance to some anxiety-inducing content. It is proposed that 'incubation of threat' is a central feature of worrying. A specially devised methodology has been used in an attempt to validate empirically some of these notions. The process of involvement which results in viewing a threat with growing apprehension has been demonstrated through vicarious worrying. Certain psychodynamic implications of the definition and of the findings are discussed.

Considering its widespread manifestations, worrying is a much neglected psychological phenomenon. Let us mention just a few of the possible factors that have brought this situation about.

As a scientific concept worrying has never been clearly defined, and to many psychologists it overlaps with the broader concept of anxiety. The centrality and diffuseness of the anxiety construct have made it difficult for a related concept to crystallize.

As an object of study, worrying is a rather elusive phenomenon. What brings it about? Under what circumstances would a person 'worry' rather than be afraid or anxious?

How should the measurement of worrying be approached? A frightened or anxious person may at least give us some physiological clues, which may be entirely absent in the case of a person comfortably reclining on a sofa and worrying himself to death about some future event. Worrying is chiefly a *cognitive process*. Such a cognitive process, even when highly intensive, retains its privacy. This is the biggest single obstacle in the way of its analysis.

Consulting the few instances in which psychologists attempted to deal—mostly as a secondary issue—with worrying, revealed that their approaches have been based upon two radically contradictory points of view:

Worrying as reality testing. In his studies of behaviour under stress, Janis has come to regard it as a necessary step to normal coping with threatening situations (1958). This notion is well stated in Janis & Leventhal (1965): '. . . there is a "work of worrying" which enables the person to cope more effectively in the long run with a painful reality situation' (p. 1369). This 'work of worrying' is assumed to begin as soon as one becomes aware of an impending danger that is perceived as potentially threatening to the self. A person who remains unworried cannot suddenly stand against the fright when the danger approaches. 'Thus, the work of worrying is conceived as a form of inner preparation that increases the level of tolerance for subsequent threat or danger stimuli. The more thorough the work of worrying, the more reality-tested the person's self-delivered reassurances are likely to be, and hence the more emotional control he will maintain under conditions of subsequent danger or deprivation' (p. 1371).

A similar notion was expressed by Arnold (1960). According to her, worrying prepares for action. It is only when the danger becomes too great, and a person is helpless, that his worrying is transformed into fear.

The functionality of worrying and anxiousness (contrasted with anxiety) was likewise stated by Odier (1956).

Worrying as a neurotic reaction. Wolfenstein (1957) does not differentiate between worrying and anxiety. More specifically, worrying is a projection of anxiety upon some external danger: 'The individual's own impulses or fantasies of punishing agencies are projected on the external world' (p. 7). The symptomatic nature of worrying is best observed when the objective threats are remote: 'Thus, in relation to remote threats, we may say that those who are relatively free from inner strain will not be likely to worry about them, and such worry when it occurs will usually indicate some emotional disturbance' (p. 8).

Interestingly, the wisdom of language points to the above-mentioned aspects of worrying. They are well represented in the various usages of the English word *care*. According to the *Concise Oxford Dictionary*, while *having cares* is equal to having worries, and *to care about* designates personal involvement, *to take care of* clearly denotes the active coping with a problem. The word *worry* itself denotes both uneasiness and anxiety as well as 'to worry a problem out', meaning to tackle it again till it is solved. Thus the cognitive, reality-testing factor, as well as the anxiety factor, should be taken into consideration in any attempt to present a comprehensive analysis of worrying. Both probably concentrate on different aspects of a complex phenomenon.

At this stage, adopting Heider's (1958) recommendation of a naive phenomenological analysis may be of some help. In contrast to both fear and anxiety, worrying appears to be mainly an active cognitive process rather than a passive experience of emotion. The worrying person is cognitively preoccupied with a problem. He keeps it in the focus of his attention and tries to solve it. Lack of external feedback turns these attempts into vicarious problem-solving. There is a constant appraisal of alternatives, their probability of occurrence, their possible consequences, etc.

In spite of its cognitive nature, the process of worrying demonstrates its relevance to anxiety. This may reveal itself in such symptoms as the consideration of unrealistic alternatives, inability to distract one's mind from the problem, magical reasoning, or obsessive repetitions of the same thoughts. Turning once again to the analysis of language we find that in addition to the already mentioned connotations of *worry*, it is also used to denote: 'Tease, harass, be continuously or intermittently troublesome, to allow no rest or peace of mind, to bite, shake or pull about with the teeth (of dogs, rat, sheep), etc.' (*Concise Oxford Dictionary*). Likewise, the French word *tracasser* means 'to be worried', as well as 'to torment somebody, to walk to and fro, to waste time, and to walk but to stay on one place' (Littré's *Dictionnaire de la Langue française*). To use the analytical framework of Miller *et al.* (1960), worrying is a case of TOTE (Test–Operate–Test–Exit) without the exit. This makes it a loop of the kind TO–TO–TO, etc. (Test–Operate–Test–Operate, etc.).

The cognitive preoccupation with a problem is a relatively time-consuming process. Once again the difference from anxiety and fear is obvious. It takes seconds to be emotionally aroused, even up to high intensities. Worrying, however, takes much longer. Analysis of a problem from various points of view, checking the various alternatives, their probability, their possible outcomes, the implications of various outcomes, etc., may take minutes, hours or, indeed, whole years. This makes it necessary to discuss the *process of worrying* rather than a *state of worrying*.

OPERATIONAL DEFINITION OF THE PROCESS OF WORRYING

The fact that both the active coping with objective threat and the anxiety factor appear to be relevant aspects of worrying does not imply that it consists of a compromise, or a mixture of the two. On the contrary, according to the position taken in this paper, worrying is viewed as depending on a very particular interaction between the two—that of *thematic synchronization*. The formal operational definition follows: *If an external threat concerning a possible future event touches upon an anxiety-inducing content, the process of worrying starts.*

External threat alone usually produces the effect of objective fear, and is not a sufficient condition for worrying; nor is anxiety alone able to produce it. What is needed is the relevance on some psychological dimension of the threat to the anxiety which then serves as an amplifier of the objective threat. The predominantly cognitive characteristic of worrying depends, however, upon the uncertainty of the future. The person must not abandon all hope. The very fact that fear and hope are mixed makes the threat a problem to be analysed. What would happen? What can be done to prepare oneself for the worst, while at the same time hoping for the best? If the approaching danger is sure to materialize, there is nothing to be done about it, but wait for it to happen. The cognitive elaboration of the possibilities which are inherent in the situation is therefore a function of the level of uncertainty.

The special position of worrying *vis-à-vis* fear and anxiety is schematically presented in Table 1.

Table 1. *Comparison of worrying with fear and anxiety*

Facets of comparison	Fear	Worrying	Anxiety
Source	External	External relevant to internal	Internal
Cognitive functioning	Minimal, mainly cognitive appraisal	Central, degrees of planning and rigidity	Minimal, mainly rigidity
Emotion	Central	Secondary	Central

THE PROCESS OF INVOLVEMENT

We have already stipulated that worrying is a process of cognitive preoccupation with a problem. Following the operational definition used in this study, we may now add that it is preoccupation with a threat; not just any threat, but one which is related to an anxiety-inducing content. The worrying person starts to think of dangers, tests his possible responses, evaluates the chances of various happenings, etc. Because of its relevance to anxiety, he is paying more attention to the threat than is its due, he is agitated beyond a realistic level, and cannot leave the problem. It is suggested that these behaviours be labelled the *process of involvement*. The person becomes gradually more and more involved in the threatening future, to the exclusion of everything else. He is, in a way, absorbed by it, and his attention is fixated.

The uncertainty of the danger causes the person to be constantly on the lookout to make an evaluation of the situation and the probabilities of various outcomes. This leads to a typical instance of 'hypervigilance' (Janis, 1962). The continuous 'cognitive appraisal' (Lazarus, 1966) of warning signals in the environment keeps the threat in the focus of attention, and further contributes to the process of involvement. Finally, uncertainty is stressful in its own right. In instances of disasters some of the strongest stress reactions were found in people who experienced uncertainty about the nature of the disaster and its effects (Baker & Chapman, 1962).

The fact that the threatening future is coming closer and closer with time is another major reason for the process of involvement. As the objective danger comes closer, the person is less able to ignore it. This should be a constant reminder that worrying is not just a 'projection of anxiety', and is intimately related to an external threat.

We may postulate that the longer the process of worrying, the more advanced is the process of involvement. While, at the start, the person's thoughts may be closely related to the objective dangers in front of him, through involvement the person enters more and more into the realm of imagination, low probabilities, agitation and helplessness. He is gradually wasting all of his 'self-delivered reassurances' (Janis & Leventhal, 1965), and becomes more susceptible to anxiety. Thus the longer the duration between the onset of the particular threat and its execution the stronger will the reaction to it become. This effect, which has been termed '*incubation of threat*' (Breznitz, 1967), has already been demonstrated in situations of unavoidable frightening events (Breznitz, 1965, 1967), and in a situation of conflicting expectations, which is more closely related to worrying (Breznitz, 1968). It is the active process of involvement, rather than time as such, which is the key factor in 'incubation of threat'. The cognitive process takes time and consequently time serves as an opportunity for this internal process to develop. It should be viewed as a necessary but not sufficient condition for 'incubation of threat', since in case of successful distraction of one's mind time would probably be irrelevant.

From all these considerations it follows that while initially the process of worrying is closer to the objective fears related to the external threat, through the process of involvement it gradually approaches closer to anxiety.

The position on this continuum depends on factors such as the intensity of the anxiety relevant to the threat, and the degree of relevance between them. The greater the anxiety, or the closer it is to the external threat, the more rapid would the advance of the process be, and the closer would it be to the anxiety pole of the dimension.

The extremes of this fear–anxiety dimension reflect to some degree the controversy in psychological literature. The fear pole is related to what Janis and others had in mind when discussing the work of worrying, while the anxiety pole is closely related to Wolfenstein's notion of worrying. The important thing to remember is that *worrying is neither of the two, but rather the process which transforms the one into the other*.

How could these theoretical speculations be submitted to a rigorous test? The process of involvement is, after all, mainly internal, and well concealed from the objective student of worrying. The following empirical study had, therefore, two major aims: first, it attempted to evaluate a particular methodological approach to the problem of worrying, and secondly, it used that methodology to test the central hypothesis of 'incubation of threat'.

METHOD

The 'social intelligence test' method. Even if we had been successful in obtaining reports of the thoughts of worrying people, there is little doubt that the reliability of such self-reported data would have been problematic. How then are we to get at these entirely subjective events? One possible approach is to make subjects guess the worrying process of other people. They are not required to respond to the other's worrying; that would be a case of vicarious instigation (Berger, 1962), but they do have to put themselves in their place and guess their thoughts. Since they have very little information to utilize, the chances are that without being explicitly asked to do so, they will project their own thoughts and feelings. The task in this study was presented as a social intelligence test. The answers to the questions were said to be available, and subjects had to guess the correct response. They were told that on the basis of their performance their social intelligence would be computed.

The task. Subjects were presented with the following problem: '*A person is waiting for another person to come. The person who is late is very dear to the one who is waiting.* On the following page you will find a list of possible thoughts that may pass through the mind of the one who is waiting. Try to guess which of the thoughts are most appropriate, and write their code number in the order that they enter his mind. You can use as many thoughts as you wish, any that you wish, and you can use any of them as many times as you see fit. Your time is unlimited'.

On the following page was a list of 14 possible thoughts with numbers attached to them. The numbers were randomly ascribed, and so was the position of each thought on the page. There were two different sets of distributions, and none had any symmetrical pattern, so as to minimize order effects of readings and any other factor of perceptual screening. At the bottom of the page was a place provided for the answer, with a starting point so as to ensure the directionality of the response sequence.

List of possible thoughts. The list included four categories of responses which were selected in a series of pretests with open-ended questions. The four categories included thoughts that: the other will come soon, the other can't come, the other doesn't want to come, and aggression towards the other. In addition, one item of self-blame and another of self-reassurance were also incorporated into the list. The items in the order of one of their random sequences were: (1) 'He will surely come soon', (2) 'Perhaps I made a mistake about the time?' (3) 'Never mind, I shall wait', (4) 'Perhaps he had a road accident', (5) 'Perhaps something happened to him', (6) 'He missed the bus', (7) 'He forgot to come', (8) 'One can after all give him up', (9) 'I will show him', (10) 'He forgot me', (11) 'He changed his mind', (12) 'This is making me angry', (13) 'There is nothing to worry about', (14) 'How the time is passing'. Items 2, 13 and 14 did not fit into any of the preconceived categories and were included mainly in order to achieve greater verisimilitude of the task.

Subjects and procedure. The subjects were 112 students of the Hebrew University attending an introductory course in psychology, and they were tested collectively. Those who finished the test moved on to the next page, which included a few questions about the task.

RESULTS

Dimensions of responding. On the basis of Guttman's scalogram analysis (Guttman, 1944) the four categories of responses were tested for scalability. In each category the items were ordered *a priori* according to the intensity with which they express the underlying dimension. As it happened, there orders agreed with those obtained by endorsement frequencies. On three categories the coefficient of reproducibility was well above the requirements, whereas in the category of items dealing with the optimistic thoughts that the other will come soon, the order of items which formed a Guttman scale was the opposite of the expected one, implying the necessity of reversing the name of the dimension. The four dimensions and their components appear in Table 2.

Table 2. *The Guttman scales and their coefficients of reproducibility*

Item (starting with top intensity)	Name of scale	Coefficient of reproducibility
4. 'Perhaps he had a road accident' 5. 'Perhaps something happened to him'	Accident scale	99%
11. 'He changed his mind' 10. 'He forgot me' 7. 'He forgot to come'	Rejection scale	97%
8. 'One can after all give him up' 9. 'I will show him' 12. 'This is making me angry'	Aggression scale	96%
6. 'He missed the bus' 3. 'Never mind, I shall wait' 1. 'He will surely come soon'	Diminishing optimism scale	93%

The unidimensionality of the scales implies that if a subject uses a high intensity item, he would use all the less intense items from the same dimension as well. If he checks item 11 ('He changed his mind'), for example, he would also check the other two items of the *Rejection* scale, i.e. 10 ('He forgot me'), and 7 ('He forgot to come'). By counting the number of items chosen from a particular scale, one arrives at the scale score which enables the reconstruction of all the answers. The fact that these categories are unidimensional scales has not only strictly methodological implications. In order to obtain this degree of order the various dimensions must be somehow represented within the subjects and to a large degree control their responding.

'Incubation of threat'. As all the four scales have at their peak (high intensity point) items which are the most negative ones (more rejection, greater accident, more aggression, and less optimism), the hypothesis of 'incubation of threat' can now be operationally stated as follows: *Within each scale, the high intensity items should appear after the low intensity items.* In other words, as the process of involvement goes on, subjects should advance along the various dimensions to more pessimistic thoughts.

Starting with the Accident scale, we find that of 41 who checked both items 4 and 5, 37 put 5 before 4. A sign test indicates that this is a significant difference ($P < 0{\cdot}001$). Since item 4 is the more threatening of the two, this supports the hypothesis. Proceeding to the *Rejection* scale we should predict the order 7 → 10 → 11. Of the 23 subjects who checked both items 7 and 10, 17 put 7 as first ($P < 0{\cdot}05$). Of the 18 who used both items 10 and 11, 11 put item 10 first ($P < 0{\cdot}05$). The hypothesis of 'incubation of threat' was therefore supported in the Rejection scale as well. Within the Aggression scale the expected sequence is: 12 → 9 → 8. Out of 33 who used both items 12 and 9, 22 put 12 before 9 ($P < 0{\cdot}05$). Out of 20 who used items 9 and 8, 17 put 9 first ($P < 0{\cdot}001$). This again supports the hypothesis. Lastly, we shall analyse the Diminishing Optimism scale. The predicted sequence is: 1 → 3 → 6. Out of 57 subjects who checked both items 1 and 3, 47 put 1 before 3 ($P < 0{\cdot}001$). Out of 33 using both items 3 and 6, 21 put 3 as first ($P < 0{\cdot}05$).

The 'incubation of threat' has thus been demonstrated in all of the four dimensions of responding. As the process of worrying develops in time, the subjects become more and more involved in the task, and choose items with intensively more negative implications. One should view this successful demonstration not only in terms of the proposed theoretical framework, but as a support of the particular methodology used. It is no small wonder that subjects in the set of a 'social intelligence test' became involved in the task to such a degree that their responses revealed the process of 'incubation of threat'.

Interdimensional dynamics. All of the analyses done so far were dealing with processes within a dimension. The data at our disposal make it possible to test also the transformations *between dimensions*. The matrix of transition probabilities between all the items revealed that the obtained data could not possibly derive from any sort of Markov chain process. Since there was an interesting indication that items of the Aggression scale have greater probability of occurring after items of Rejection than of following any other items, it was thought worthwhile to pursue this point further. It was found that 77 per cent out of all the instances in which the most intense Aggression (item 8) was used followed items of the Rejection scale. While 31 per cent of the Rejection scale are followed by Aggression scale items, only 3 per cent of Diminishing Optimism scale items, and 4 per cent of Accident scale items do so ($P < 0{\cdot}001$). These results may indicate the psychological importance of *reciprocity* (or 'mutuality' to use a term of Tagiuri *et al.*, 1958). Thoughts of rejection by the other provoke anger, since the other deserves punishment. By disliking him the rejection becomes mutual.

In order to obtain additional information about the relationships between the various scales, the intercorrelations between all of the scale scores were computed. They are presented in Table 3.

The particular order of the variables in Table 3 fulfils the requirements of a perfect simplex (Guttman, 1954), which indicates that they combine into a single master dimension. Whereas the exact nature of this dimension is still unknown, its interesting feature is that Aggression takes its position between Diminishing Optimism and Rejection. In conjunction with the previous analysis, it may suggest that aggression may not only fulfil the tendency to reciprocate rejection, but that it could serve as an optimistic thought in disguise. If one thinks of how he is going to teach the other a lesson, this implies that the other is after all going to come, and that nothing serious happened to him or to their relationship. The apprehensions concerning the other may thus be transformed into anger, to one's psychological advantage.

Table 3. *Intercorrelations between scale scores*

	Accident	Rejection	Aggression	Diminishing optimism
Accident scale	1·00	0·45	0·36	0·14
Rejection scale	0·45	1·00	0·61	0·38
Aggression scale	0·36	0·61	1·00	0·45
Diminishing optimism scale	0·14	0·38	0·45	1·00

DISCUSSION

Although the validity of the social intelligence test method in the study of worrying is difficult to ascertain, its initial performance is rather promising. The assumption that subjects have only minimal information about other people's thoughts in such situations, and consequently cannot in their responses rely on stereotypes, appears to be justified, since 50 per cent of them without being explicitly asked about it stated that they were 'guessing the correct answer' on the basis of their own experience. Such a confession of projection was implicit in most of the remaining subjects' reactions to the task, so that we may safely take it for granted. As a matter of fact, people may be actually interested in demonstrating that their worries resemble the modal behaviour of others in similar situations, so that the 'social desirability' factor (Edwards, 1957) would not be a major source of unreliability.

While many reported the task to be on the whole relatively easy because of their frequent experience of such a situation, they viewed the ordering of the items as the major source of difficulty. Assuming that a person projects his private thoughts, this difficulty is understandable, since the task of ordering the items interferes with the spontaneous 'train of thought'. While the open-ended system used in the pretest did not suffer from this, the advantages gained by designing the items more than outweighed this consideration. The significant results concerning the central hypothesis tested lend, of course, additional empirical justification to the methodology.

The results, however, may have only a limited applicability to other types of worrying situations, and to generalize from this study would be unjustified. A whole set of different situations ought yet to be tested with this method in order to warrant bolder generalizations. Even the relatively simple problem that served as the stimulus in this study was interpreted differently by different individuals. They varied in their interpretation of the setting, the sexes of the participants, the duration of waiting, etc.

Another theoretically interesting source of individual differences may be the particular focus of anxiety in a given standard situation. Thus, in our case, the major source of anxiety could be the potential loss of a love object (Accident), or the potential loss of love (Rejection). Whether such relevant anxiety should reveal itself in greater preoccupation with the relevant dimension, or, rather, in the denial of such dangerous thoughts is a question to which the answer is still missing even within the broader context of personality testing. Whatever their interpretation, some interesting individual differences in this area were indeed found in this study. As an example, it may be mentioned that 13 subjects used the strongest item of the Accident scale (item 4), without using at all even the weakest item of the Rejection scale (item 7). It is possible that the rejection anxiety of some was so intense that they preferred to think of a road accident rather than to think that the other forgot to come. Further research may attack this problem directly by the independent measurement of these anxieties.

The analysis of worrying and the preliminary study reported in this paper reveal the complexity and potential richness of this area of investigation. While the methods are difficult, the questions many and complicated, their relevance to general psychology must not be overlooked.

The author wishes to thank Mrs Eshel and Miss Bar-Hilel for their generous help at various stages of this study.

REFERENCES

ARNOLD, M. B. (1960). *Emotion and Personality*. New York: Columbia University Press.

BAKER, G. W. & CHAPMAN, D. W. (eds.) (1962). *Man and Society in Disaster*. New York: Basic Books.

BERGER, S. M. (1962). Conditioning through vicarious instigation. *Psychol. Rev.* **69**, 450–466.

BREZNITZ, S. (1965). The influence of the time interval between a threat with a frightening stimulus and its execution upon fear reaction in humans. (Doctoral dissertation, The Hebrew University.)

BREZNITZ, S. (1967). Incubation of threat: duration of anticipation and false alarm as determinants of the fear reaction to an unavoidable frightening event. *J. exp. Res. Person.* **2**, 173–179.

BREZNITZ, S. (1968). 'Incubation of threat' in a situation of conflicting expectations. *Psychol. Rep.* **22**, 755–756.

EDWARDS, A. L. (1957). *The Social Desirability Variable in Personality Assessment and Research*. New York: Dryden Press.

GUTTMAN, L. (1944). A basis for scaling qualitative data. *Am. sociol. Rev.* **9**, 139–150.

GUTTMAN, L. (1954). A new approach to factor analysis: the Radex. In P. F. Lazarsfeld (ed.), *Mathematical Thinking in the Social Sciences*. Glencoe: Ill.: Free Press.

HEIDER, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.

JANIS, I. L. (1958). *Psychological Stress*. New York: Wiley.

JANIS, I. L. (1962). Psychological effects of warnings. In G. W. Baker & D. W. Chapman (eds.), *Man and Society in Disaster*. New York: Basic Books.

JANIS, I. L. & LEVENTHAL, H. (1965). Psychological aspects of physical illness and hospital care. In B. B. Wolman (ed.), *Handbook of Clinical Psychology*. New York: McGraw-Hill.

LAZARUS, R. S. (1966). *Psychological Stress and the Coping Process*. New York: McGraw-Hill.

MILLER, G. A., GALANTER, E. & PRIBRAM, K. H. (1960). *Plans and the Structure of Behavior*. New York: Holt.

ODIER, C. (1956). *Anxiety and Magic Thinking*. New York: International Universities Press.

TAGIURI, R., BRUNER, J. S. & BLAKE, R. R. (1958). On the relation between feelings and perception of feelings among members of small groups. In E. E. Maccoby, T. M. Newcomb & E. L. Hartley (eds.), *Readings in Social Psychology*, 3rd ed. New York: Holt.

WOLFENSTEIN, M. (1957). *Disaster: A Psychological Essay*. Glencoe, Ill.: Free Press.

Manuscript received 11 June 1970

Revised manuscript received 28 October 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 280–281
Printed in Great Britain

Patterns of Evaluation to R.P., South Welsh and Somerset Accented Speech

BY HOWARD GILES
University of Bristol

Lambert (1967) has shown in many cultures that regional dialect is a significant cue in assessing personality from voices. Similar research in Britain, however, has been limited to two studies. Strongman & Woosley (1967) studied the reactions of northern and southern English listeners to the male 'matched-guises' of Yorkshire and London accents. Although no group differences emerged, judges agreed that the London speakers were more self-confident, while the Yorkshire voices seemed more honest and reliable. The northern judges themselves perceived their own group as significantly more generous, goodnatured and kindhearted. Cheyne (1970), studying evaluations to Scottish and English voices, found that both Scottish and English listeners rated male English speakers as possessing more leadership, intelligence, ambition and self-confidence. The Scottish listeners, like the northern judges, were also *accent loyal*, evaluating their own group as more generous, good-hearted, friendly, humorous and likeable, with the English judges also conceding friendliness to the Scottish voices.

Assuming that the loosely-termed 'London' and 'English' accents in the above studies may be regarded as representative of southern Standard English or Received Pronunciation (R.P.), then a trend in evaluation may be discerned from these results. Lambert (1967) was able to group the majority of his personality traits 'into three logically distinct categories of personality', namely *competence* (intelligence, self-confidence, etc.), *personal integrity* (reliability, kindness, etc.) and *social attractiveness* (sociability, sense of humour, etc.). Thus at least with regard to male voices, it would seem that speakers of R.P. may attract stereotyped personality impressions of greater competence from listeners than speakers of 'non-standard' regional accents—a phenomenon even accent loyalty tends not to deny. However, regional accented judges, and to a lesser extent R.P. judges also, seem to consider 'non-standard' speakers as possessing greater personal integrity and social attractiveness in relation to R.P. speakers. The present study was designed to determine whether this same evaluative trend would also be reflected in a personality assessment from two other regional accents relative to R.P.—South Welsh and Somerset. In a recent study, Giles (1970) found that regional accents could be considered as occupying relative positions along an evaluative continuum ranging from high to low status. Since R.P., South Welsh and Somerset represented high, intermediate and low positions respectively, it was hypothesized that a similar pattern of reactions would emerge in the competence personality traits, with perhaps South Welsh still attaining more favourable ratings over Somerset in the personal integrity and social attractiveness traits.

Tape-recordings were made of two male speakers reading a standard 2-min., neutral passage of prose quite realistically in each of three accents: R.P., South Welsh and Somerset. The speakers attempted to assume the same paralinguistic features and personality throughout the recordings, with the voices being presented to the judges in a balanced order of the two speakers.

Ninety-six subjects were randomly selected according to a 2 × 2 matrix of region (South Wales and Somerset) and age, affording group means of 12 years 5 months and 17 years 6 months. The four subject groups were matched for sex, comprising 12 males and females each. The evaluative task was explained to the subjects in terms of forming immediate impressions of personality from voices, with the notion of 'accent' not being mentioned. The subjects were required to rate the six voices on 18 bipolar traits (those used by Strongman & Woosley, 1967) with seven-point scales, the order of trait presentation being reversed for each alternate voice.

The average of each subject's ratings for each of the three accents for every personality scale was computed, thus providing the raw data for analyses of variance on each trait.

A summary of the results is given in Table 1, where overall age and regional differences in evaluation emerged. It was found that on five traits, four of which were competence traits,

Table 1. *Summary of significant differences from the analyses of variance*

Age differences	Regional differences	'Standard' rated most favourably	Regionals rated more favourably
Good looks*	Generosity*	Good looks*	Seriousness*
Ambition*	Good looks*	Ambition*	Talkativeness*
Intelligence*	Seriousness**	Intelligence*	Goodnaturedness*
Self-confidence**	Goodnaturedness*	Self-confidence*	Humorousness*
Determination**	Honesty*	Determination*	
	Imagination**	Industriousness*	
	Ambition*		
	Intelligence**		
	Kindheartedness*		
	Industriousness*		

* $P<0{\cdot}01$; ** $P<0{\cdot}05$.

17-year-olds rated the 'standard' accent more favourably but the two regionals less favourably than the younger group. On 10 traits, the Somerset sample evaluated all three accents less favourably than the Welsh listeners.

Also in Table 1, it can be seen that on six traits (five of which were competence scales) R.P. was rated significantly more favourably than either regionals. Indeed, on all five of these competence traits the pattern of relative accent prestige values was reflected in the evaluative trends. Moreover, on four traits concerning personal integrity and social attractiveness, the regional accents were evaluated more favourably than the 'standard', yet only the trait of talkativeness reflected the hypothesized more favourable ratings for the South Welsh speakers in relation to the Somerset voices. In contrast, as regards goodnaturedness, the *Somerset* voices were most favourably perceived, and so a consistent patterning of relative accent prestige values seems to be maintained only in the evaluation of the competence traits of personality. It would seem that in Britain personal integrity and social attractiveness may not be so easily subjected to evaluative norms of social prestige, but may perhaps be dependent on the more diverse, yet nonetheless stereotyped, social qualities and temperaments characteristic of specific regional communities. On the other hand, with the traits of good looks and humorousness there was a tendency (significant in the latter) towards *accent loyalty*. In other words, judges would perceive voices representative of their own speech community in particular respects more favourably than the other regional accent presented.

It would appear then that the more prestigious 'standard' accented voice in Britain is stereotyped as commensurate with highly favourable personality traits of competence—a finding seemingly independent of listeners' regional membership. Nevertheless, 'non-standard' regional accented speakers would seem to be stereotyped more favourably than R.P. speakers with respect to personal integrity and social attractiveness. However, it would seem that competence personality impressions of accented speakers are to a certain extent a function of relative accent prestige values and social maturity, with regional membership displaying important evaluative orientations across most personality traits.

I would like to thank Mr P. F. Powesland, Bristol University, for his excellent voice productions and supervision of the project.

REFERENCES

CHEYNE, W. M. (1970). Stereotyped reactions to speakers with Scottish and English regional accents. *Br. J. soc. clin. Psychol.* **9**, 77–79.
GILES, H. (1970). Evaluative reactions to accents. *Educ. Rev.* **22**, 211–227.
LAMBERT, W. E. (1967). A social psychology of bilingualism. *J. soc. Issues* **23**, 91–109.
STRONGMAN, K. T. & WOOSLEY, J. (1967). Stereotyped reactions to regional accents. *Br. J. soc. clin. Psychol.* **6**, 164–167.

Manuscript received 26 July 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 282–283
Printed in Great Britain

Test–Retest Results on the Conservatism Scale Completed under Conditions of Anonymity and Identification

BY D. K. B. NIAS, G. D. WILSON
Institute of Psychiatry, University of London
AND JOAN M. WOODBRIDGE
Bath College of Education

The C scale has been reported to have a very high (0·94) split-half internal consistency (Wilson & Patterson, 1968), in spite of the wide range of social attitudes covered by the 50 items. Data on the factor structure of the scale and evidence for its validity are summarized by Wilson & Patterson (1970). The present study was carried out to provide information, firstly on test–retest reliability, and secondly on the effects of the scale being completed anonymously or by name.

Four groups of female students from a college of education were tested twice, with an interval of 12 weeks. Group 1 were asked on the first occasion of testing to give their names, and on the second occasion to remain anonymous. (The two sets of answer sheets were later linked by employing a code system.) For group 2 the procedure was reversed; group 3 were asked to give their names on both occasions and group 4 to remain anonymous. Table 1 shows the mean C scores for each group on the two occasions of testing. Reliabilities were computed using product-moment correlations.

Table 1. *Means, standard deviations and reliabilities for the four groups*

Group	*n*	Mean (and S.D.)		Reliability
		Test	Retest	
1. Named-Anonymous	28	45·1* (12·2)	42·4 (13·4)	0·94
2. Anonymous-Named	25	42·2 (11·2)	41·3 (10·6)	0·82
3. Named	15	39·3 (10·9)	38·3 (11·4)	0·84
4. Anonymous	14	44·8 (12·7)	43·5 (14·0)	0·90
Total	82	43·1 (11·8)	41·5 (12·2)	0·89

* C scale scores range from 0 to 100, with high scores indicating 'conservatism'.

The test–retest reliability for the total sample of 82 was 0·89. This coefficient compares very favourably with those reported for other scales, and suggests that the straightforward item-format adopted by Wilson & Patterson has a promising future in questionnaire research. The difference between the reliabilities for the four subgroups is not significant ($\chi^2=4{\cdot}5$; d.f. $=3$). It may be noted that the C scores were slightly lower on the second occasion of testing. The overall mean change in score was 1·6, which is statistically significant ($t=2{\cdot}4$; d.f. $=81$; $P<0{\cdot}05$). This change towards 'liberalism' may be attributable, of course, to the effects of retesting or to a change in attitudes for the group over the 12 weeks.

Patterson & Wilson (1969) report that when given the option of whether or not to reveal their identity, subjects who elected to remain anonymous obtained higher C scores. This result was interpreted in terms of a hypothesized preference among 'conservative' subjects for generally 'playing it safe'. Another possible explanation, however, is in terms of subjects being prepared to admit to more conservative attitudes when completing the scale anonymously. In other words, conservative responses may be seen as less socially acceptable and so tend to receive endorsements only under conditions of anonymity.

An opportunity to test this hypothesis was provided in the present study. By combining the scores for the Named condition from groups 1 and 3 we compared the mean with that obtained for the Anonymous condition from groups 2 and 4; the two means were the same at 43·1. Another comparison was made by combining the scores for groups 1 and 2 so as to form a 'cross-over' design; the mean under the Named condition was 43·2 and under the Anonymous condition 42·3. This difference is not significant ($F = 1{\cdot}4$; d.f. = 1, 51), but it is relevant to note that its direction is consistent with Wilson & Patterson's (1969) suggestion of a positive association between 'conservatism' and 'social desirability' based on a correlation of the C scale and the L scale of the Eysenck Personality Inventory. Anyway, the present results indicate that there was little or no difference in social desirability between conservative and liberal responses for the sample we studied, and so tend to support the original contention that preference for anonymity is a reflexion of conservatism. Finally, it may be noted that our sample's mean C score of 43·1 is higher than that reported for other student groups, and probably reflects the fact that our subjects were drawn from a single-sex college.

REFERENCES

Patterson, J. R. & Wilson, G. D. (1969). Anonymity, occupation and conservatism. *J. soc. Psychol.* **78**, 263–266.

Wilson, G. D. & Patterson, J. R. (1968). A new measure of conservatism. *Br. J. soc. clin. Psychol.* **7**, 264–269.

Wilson, G. D. & Patterson, J. R. (1969). Conservatism as a predictor of humor preferences. *J. consult. clin. Psychol.* **33**, 271–274.

Wilson, G. D. & Patterson, J. R. (1970). *Manual for the Conservatism Scale.* Windsor: N.F.E.R.

Manuscript received 11 *September* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 284–285
Printed in Great Britain

Interference, Extraversion and Paired-Associate Learning

BY RONALD N. BONE

West Virginia Wesleyan College

Recently, McLaughlin & Eysenck (1967) found extraverts superior to introverts on both easy and difficult paired-associate tasks. These data were interpreted by assuming that introverts possess higher cortical arousal which interferes with the integration on the lists. This inference is supported by Jensen's (1964) finding of introverts being less resistant to response competition in a serial-learning task.

If these findings have generality, similar effects should be obtained using other verbal learning tasks, especially those containing built-in interference. Such a task has been utilized by Spence (1963) involving a paired-associate list of stimulus and response members having strong associative strength. These pairs were scrambled, however, so that practically no associative connexion existed between pairs, though each stimulus had a strong association among the response members elsewhere in the list. Such a list was viewed as comparable to the A–B, A–B_r transfer paradigm (stimuli and responses of the first list are re-paired to form a second list), assuming that the first list (A–B) is composed of strongly associated word pairs brought into the laboratory via natural language habits. The second list (A–B_r) was composed of a scrambled version of the first list. When compared to a control list consisting of no cross-associates, she found performance to be significantly inferior. Presumably this effect is due to both forward (S–R) and backward (R–S) interference (McGovern, 1964).

Using Spence's (1963) version of the A–B_r paradigm, it was predicted that introverts would display more errors than extraverts on a list containing re-paired primary associates. An equal number were predicted for introverts and extraverts on a list containing unrelated words.

The subjects were introductory psychology students participating as part of a class requirement. Two weeks prior to the experiment 127 students were given the Maudsley Personality Inventory (Eysenck, 1959). Twenty subjects scoring in the introverted direction and 20 in the extraverted direction were chosen to represent extreme groups. These groups were, in turn, randomly divided into two groups of 10 each to learn either interference or control lists. Mean extraversion scores were 40·37 and 14·28 for interference lists; 39·08 and 15·21 for control lists. While neuroticism was not investigated as a variable in this study, it may be pointed out that mean neuroticism scores did not differ in the four treatments.

The paired-associate lists consisted of 10 pairs, the stimuli being the same in both interference and control lists. The mean cross-associate strength for the interference list was 0·60; for the control list 0·00. Association values were obtained from Bilodeau & Howell's (1966) norms for discrete associations. Associative strength between individual stimulus members, response members and S–R pairs was kept to a minimum. All of the stimulus words were of high frequency.

Each was given standard paired-associate instructions using the anticipation method. The lists were presented in four random orders to prevent serial learning. The intertrial interval was 6 sec. Learning was taken to a criterion of one errorless trial.

The results are shown in Tables 1–3 and lend strong support to the hypothesis that introverts are more prone to interference than extraverts on tasks containing built-in interference, but not on tasks where interference is absent or at a minimum.

These findings are for the most part consistent with the study by McLaughlin & Eysenck (1967), the only difference being that they found the performance of extraverts superior on both difficult and easy lists. Possibly the difference between these two studies is more apparent than real and lies in the definition of easy and difficult tasks. McLaughlin & Eysenck (1967) defined difficulty in terms of formal response similarity, with difficult and easy lists containing formal high and low response similarity respectively. The easy list could still be difficult since nonsense syllables of low formal similarity may place considerable demands on the response learning phase of paired-associate learning (Underwood *et al.*, 1959).

Table 1. *Mean trials to criterion and mean number of errors*

Treatments	Trials to criterion	Errors
Extravert		
Interference	15·20	55·00
Control	8·90	21·90
Introvert		
Interference	18·20	77·90
Control	8·40	20·50

Table 2. *Analysis of variance for the trials to criterion*

Source of variation	D.F.	M.S.	*F*	*P*
Extraversion	1	15·62	1·69	
List (B)	1	648·02	69·98	<0·001
A × B interaction	1	30·63	3·31	
Error	36	9·26		

Table 3. *Analysis of variance for the number of errors*

Source of variation	D.F.	M.S.	*F*	*P*
Extraversion (A)	1	1152·62	4·72	<0·05
List (B)	1	20475·62	83·80	<0·001
A × B interaction	1	1476·23	6·04	<0·05
Error	36	244·34		

In conclusion, it is suggested that verbal learning tasks are ideally suited for the investigation of personality variables. However, anyone attempting to work in this area should pay particular attention to an excellent article by Goulet (1968) concerning methodology.

REFERENCES

Bilodeau, E. A. & Howell, D. C. (1966). Free association norms: discrete and continuous methods. (Office of Naval Research Department of the Navy, Washington, D.C. Tulane University. Technical Report no. 1 for contract Nonr-475 (10).)

Eysenck, H. J. (1959). *Manual of the Maudsley Personality Inventory*. London: University of London Press.

Goulet, L. R. (1968). Anxiety (drive) and verbal learning: implications for research and some methodological considerations. *Psychol. Bull.* **69**, 235–247.

Jensen, A. R. (1964). Individual differences in learning: interference factor. (Co-operative Research Project no. 1867. Office of Education, U.S. Department of Health, Education and Welfare.)

McGovern, J. B. (1964). Extinction of associations in four transfer paradigms. *Psychol. Monogr.* **78**, no. 16. (Whole no. 593.)

McLaughlin, R. J. & Eysenck, H. J. (1967). Extraversion, neuroticism and paired associate learning. *J. exp. Res. Person.* **2**, 128–132.

Spence, J. T. (1963). Associative interference on paired-associate lists from extra-experimental learning. *J. verb. Learn. verb. Behav.* **2**, 329–338.

Underwood, B. J., Runquist, W. N. & Schulz, R. W. (1959). Response learning in paired-associate lists as a function of intra-list similarity. *J. exp. Psychol.* **58**, 70–78.

Manuscript received 28 *May* 1970

Revised manuscript received 5 *October* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 286–288
Printed in Great Britain

A Fully Automated Intellectual Test

BY HARRY BRIERLEY

Professorial Unit of Psychological Medicine,
General Hospital, Newcastle upon Tyne

A number of writers, e.g. Miller (1968) and Peck & Gathercole (1968), have commented upon the use of automated testing techniques. Whilst methods such as those of Heron (1953) and Zenkteler (1968) represent efforts to make personality inventories more palatable and to simplify scoring, little evidence is available concerning the use of automated testing of intellectual functioning. Gedye (1966) has reported the use of an automated technique administering a novel procedure but not one readily compared with conventional intellectual testing methods. Experiments with automated forms of Perceptual Mazes and Raven's Matrices have been employed by Elithorne and Berger (personal communications, 1969).

The present paper reports the use of an automated test designed for research purposes (Brierley, 1969). In this procedure, the patient was isolated in a room with the testing apparatus, receiving no live oral assistance or instructions, and making no verbal or written responses.

The nature of the proposed research demanded a test procedure where precise record of response time could be obtained, where only one item at any time was available to the subject, and where the administration was carefully controlled.

The test used as a model was Anstey's Dominoes (Vernon & Parry, 1949). This test was designed as a parallel test to Raven's Matrices with a reliability of 0·9 and was found to have a rather higher *g* saturation than the Matrices. On the basis of the original Dominoes Test, two sets of 48 items were devised. All these items used the 3 × 3 domino matrix in which the bottom right-hand domino was the blank answer domino. A practice set of eight other items was also designed.

These items were presented on a 3 ft. × 2 ft. 6 in. display panel at a distance of 8 ft. from the subject, who recorded responses on an answer panel of 14 micro-switches. Response switches were numbered 0 to 12, corresponding to the possible test item answers, i.e. number of dots expected on the answer domino; the remaining switch was used to signify abandonment. Standard instructions were administered by tape-recorder to the subject, who was seated alone in a sound-insulated room. The responses were recorded on a pen-recorder, which gave times of response and distinguished right, wrong, and abandoned answers. Each response cancelled the item and presented the next item after an interval of 3 sec.

The circuit design was based on banks of 'latching' relays scanned by a uniselector mechanism. This was operated by a Teletype five-wire punched tape reader, which also set up the answer checking mechanism. A full description of the circuitry is available in Brierley (1969) and a circuit diagram can be obtained from the author.

Two initial tests each of 48 items were presented to a group of normal subjects and an item analysis carried out. This led to two tests of 24 items designed in cycles of eight items, the overall difficulty of the cycles being very similar but the item difficulties within each increasing. The first test was preceded by the eight practice items fully explained by the tape-recording. The retest was also preceded by these items but without explanation.

A group of 40 subjects who were mostly members of hospital staff, between 15 and 40 years of age, and above average intelligence, were given both tests. These subjects also performed part 1 of the AH4 test. In almost all cases, the retest was carried out within a week of the first test.

Table 1 summarizes the reliability obtained from this testing in terms of number-correct scores. A test–retest reliability between the two 24-item tests of 0·86 was obtained. Using the best obtainable estimate of the standard deviation of AH4 part 1 for a general population group this reliability value is corrected for range restriction to 0·88. A split-half correlation for the two tests combined is 0·89. Corrected to full length by the Spearman–Brown formula this gives a reliability of 0·94, without correction for range restriction. This evidence suggests quite a satisfactory level of reliability for the tests, quite compatible with the reliability of the original Domino Test itself.

Table 1. *Reliability coefficients for automated dominoes tests*

	Test I	Test II	Total test
Test–retest			
Test I	—	0·86	0·97
Test II	—	—	0·96
Split-half			
Odd–even items	0·84	0·80	0·89

A comparison with the results of AH4 part 1 is shown in Table 2. Again range correction of this total test correlation with AH4 yields a value of 0·77 compared with a correlation of Raven's Matrices with AH4 part 1 of 0·66.

Table 2. *Relationships of automated dominoes scores with AH4 Part 1*

			Accuracy		
	Speed	Continuance	High difficulty	Low difficulty	Total score on both tests
AH4 Pt. 1	−0·204	0·466	0·604	0·435	0·655
Speed		−0·075	−0·146	+0·103	−0·037
Continuance			0·652	0·248	+0·592
Accuracy					
High difficulty				0·423	0·829
Low difficulty					0·453

Other values examined are those derived by Furneaux (1960) relevant to the determination of his difficulty scales. Individual speed constants (K_{Tr} values), continuance (measured as critical difficulty level ${}_{c}D_{\mathrm{tr}}$), and the accuracy of response on groups of items representing low and high difficulty were correlated with AH4 part 1 results as shown in Table 2. Speed constants estimated from the two automated domino tests show moderate reliability (0·54) but have very low intercorrelations with the other values and in particular with AH4. Even though the subjects were of better than average intelligence, performance on AH4 is more closely related to accuracy at higher difficulty and continuance than to speed as Furneaux defines it.

It would seem that from these results little if anything is lost in terms of reliability and validity in a test of the Dominoes kind as the result of total automation of the procedure when subjects are of at least average ability. At the same time this automated procedure enhanced subject motivation, even though the degree of isolation and lack of cues about the results of the test were found unpleasant by some subjects. However, total isolation of the subject from the psychologist has other problems. It gives little opportunity for observing, and sparse information about, anomalies of performance particularly important to the clinical psychologist. Although practical considerations such as time saving might be taken into consideration, it is doubtful if testing of this nature will contribute greatly to clinical practice. On the other hand, far greater experimental control and more precise measurements are practicable for research purposes.

The author gratefully acknowledges the support of the Medical Research Council in this work.

REFERENCES

Brierley, H. (1969). The intellectual speed and accuracy of psychiatric patients diagnosed as neurotic. (Unpublished Ph.D. thesis, University of Leeds.)

Furneaux, W. D. (1960). Intellectual abilities and problem-solving behaviour. In H. J. Eysenck (ed.), *Handbook of Abnormal Psychology*. London: Pitman.

GEDYE, J. L. (1966). Experience with an automated psychological testing system designed for clinical use. *Bull. Br. psychol. Soc.* **19**, no. 63, A18–19. (Abstract.)

HERON, A. (1953). Old wine in a new bottle: a new method of presenting a psychiatric questionnaire. *Bull. Br. psychol. Soc.* no. 20, p. 15. (Abstract.)

MILLER, E. (1968). A case for automated clinical testing. *Bull. Br. psychol. Soc.* **21**, 75–78.

PECK, D. F. & GATHERCOLE, C. E. (1968). Automation techniques in clinical psychology. *Bull. Br. psychol. Soc.* **21**, 161–165.

VERNON, P. E. & PARRY, J. B. (1949). *Personnel Selection in the British Forces.* London: University of London Press.

ZENKTELER, M. (1968). A partial answer to the problems of instrumentation in automation of psychological testing. *Bull. Br. psychol. Soc.* **21**, 101. (Abstract.)

Manuscript received 27 November 1970

Book Reviews

Social Psychology. By Leon Mann. Sydney: Wiley & Sons Australasia, 1970. Pp. xiii + 165. £2.50; paper, £1.25.

This book is one of a series designed to make good a deficiency experienced by Australian psychologists—'the relative lack of text and reference books written by Australians primarily for use in Australia' (p. v). Having been alerted to this somewhat parochial aim, a British reader might anticipate a book which had only limited relevance for him. But his expectation would be disconfirmed.

The author received his M.A. from Melbourne and his Ph.D. from Yale, and he has since taught at Harvard. He has written a splendid introductory book which deserves wide adoption in this country, and no doubt in other countries too. The illustrations are in fact mainly of a European or North American background. (We learn, for instance, that the rector of Lilley in Hertfordshire wrote a book in 1695 predicting that the world would end in 1697; and that in 1698 he wrote a second book complaining that this had now happened but that no one had noticed.) But it is pleasant also to learn about the characteristics of Australian football queues (p. 87), and of subway passengers in Tokyo ('where slippery coats are sold to travellers to ease their way through the sardine packed throng, and students are employed at peak hours to cram passengers into trains'; p. 17).

The author has a psychological (rather than a sociological) approach to social psychology, and he manages to combine a hard-headed laboratory perspective with a more humanistic treatment of his material. The chapter contents are fairly orthodox and comprehensive, and of particular interest is a final chapter which integrates several types of material on decision-making. Throughout the book the author is cautious about accepting results and theories at face value, and he regularly qualifies or adds to the conclusions of other writers. The writing is clear and precise, and the book is generally pitched at a level which makes it very suitable for first-year teaching. Other cheap texts for first-year courses are all rather inadequate, and certainly none can match this one for completeness, rigour and readability.

PETER WARR

Freedom and Justice Within Walls: The Bristol Prison Experiment. By F. E. Emery. London: Tavistock Publications, 1970. Pp. 119. £1.40.

This volume is a contribution to the 'Studies in Social Ecology and Pathology' series, and contains a report and an evaluation, sponsored by the Home Office, of an experimental alteration in the regime of Bristol Local Prison.

The experimental changes, known as the 'Norwich System', were introduced to Bristol during the period 1958–60 and had previously been applied to other provincial prisons. The main features of the altered regime were: (1) an extension of the period in the day when free association among prisoners could take place; (2) an attempt to foster more stable informal relationships between prison officers and inmates, mainly by allocating responsibility for certain aspects of inmate welfare to specific pairs of officers.

Dr Emery approaches the task of appraisal by means of a theoretical framework that is essentially social-psychological, drawing on such sources as Sykes and Goffman, and combining them with his own systems perspective. The purpose is the construction of a series of suggestive predictions about the outcome of the experiment, rather than the provision of a coherent and theoretically integrated explanation of the main features of prison life.

The probable effects forecast included a decrease in the general level of tension in the prison, the absence of inmate exploitation of the greater freedom by breaking regulations, and an increase in the stability of officer–inmate relations (which would, however, require more delicate management by the governor).

The main sources of data were before-and-after interviews with 30 prisoners and 20 staff, observations by means of activity sampling, and the record of reported disciplinary incidents for the 64 weeks before and after the transitional period in which the changes were introduced.

The empirical data substantiated the main predictions. Increased voluntary association was regarded by inmates as an amelioration of their living conditions and was used harmlessly for joint relaxation. Staff attitudes swung to a belief in its value, no doubt partly as a result of the decline of disciplinary incidents during association. Half of the dissenters on the staff side held reservations based on the increased duties rather than increased association *per se*. Incidents involving physical or verbal aggression during the non-work periods showed a marked decline, though this did not carry over into work shifts. Despite an increase in the inmate population over the experimental period, the indices used by Emery to measure staff–inmate tension remained stable. There were indications of the breaking of traditional stereotypes and the opinion of each group was raised in the eyes of the other.

Despite the long delay between project and publication, this research remains very relevant in a post-Mountbatten prison system. Emery's conclusions highlight once again the basic structural contradiction in prisons, namely that while from a custodial viewpoint, a degree of distance and impersonality between inmate and officer preserves the latter's authority, these characteristics frustrate a rehabilitative orientation, working, for example, through counselling techniques. Any administrative development that marries these roles in the same personnel, or perhaps even in the same institution, is bound to be self-defeating.

While Dr Emery has produced a competent and detailed analysis, introducing and monitoring the experiment in a rigorous and sophisticated manner, the extremely circumscribed nature of the brief he was given (arising from the limited scope and content of the experimental changes) raises a concern familiar to those carrying out commissioned social research. Is the sponsorship of this study a reflexion of a genuine concern to assess the values underlying the prison system, or will it be used as a placebo, a substitute for and a distraction from research which, among other things, could have politically hazardous implications? When, for instance, will the Home Office feel able to commission a study in the taboo area of inmate self-government? Is there not already ample evidence that incarceration in its traditional form is inhuman and counterproductive?

The pointer for the future that this book leaves is pessimistic. The prison system is still unable to confront the basic reality stated many years ago by Sir Alec Paterson, perhaps the greatest British prison reformer: 'it is impossible to train men for freedom in conditions of captivity'.

KENNETH LEVINE

Models of Thinking. By Frank George. London: Allen & Unwin, 1970. Pp. 116. £1.75.

While reading Frank George's book I kept recalling *Plans and the Structure of Behavior* by Miller, Galanter and Pribram. That book outlines the potentialities of cybernetic concepts and the limitations of stimulus-response ones for psychological theory. It is a modest book—discursive rather than definitive, speculative rather than didactic. It sets out a case for looking at things in a particular way.

In contrast, Professor George's book, which is about the cybernetic approach to thought, strikes me as dogmatic and pretentious. This is an enormous pity, because the author has so much knowledge at his finger-tips. (The impression is that it remains at his finger-tips.) The first part consists of a hasty summary of relevant psychological topics, and the second part of an exposition of formal models by Hovland, Hunt, Kochen and Banerji, among others.

But there is little connecting thread: snippets of information are thrown out in brisk paragraphs rather like the 'gobbets' offered for comment in Oxford examinations. For instance, in discussing theories and models, it is irrelevant to be told that 'Fitch and Barry provided a model for Hull's first theory of learning, and the Propositional Calculus is an interpretation of the same model as is interpreted as the Calculus of Classes'. ('What?') Similarly, in discussing explanation it is a bit much to quote an expression by Carnap when a few words would have sufficed. In a brief discussion of perceptual learning it is noted that one psychologist comes under two headings in one theoretical classification and under a single heading in another. Nothing of consequence seems to follow from this observation. There are mistakes. Thorndike's experiments with cats are *not* instances of classical conditioning. Experimental psychologists are transformed into behaviourists by the definition of a behaviouristic study as 'one which is objective, or that is carried out in non-introspective terms'.

The important distinction between artificial intelligence and computer simulation becomes one between 'synthesis' and 'simulation'. We are then told that the author's interests lie *between* the two! There is no mention of the central (unsolved) problem of simulation. How do we know when a computer is simulating human behaviour? And by what licence may we infer similarity of underlying mechanisms? There is not even a reference to Turing's classic (1950) paper, 'Computing Machinery and Intelligence'.

In fact, this book smacks of *scientism*. It is pervaded by jargon and the liberal use of symbolic logic which appears hardly necessary to sustain what little argument exists. There is, too, an old-fashioned ring of operationism and positivism about it. I sense the tacit assumption that computers and programs are conceived to be ends in themselves, rather than mere devices which may increase knowledge, when we know exactly what to compare and what to measure. This way of talking may lead the psychology student to believe that powerful techniques are substitutes for insight and imagination in research.

I sincerely hope this book will help someone. With luck somebody may be inspired by its vigorous flirtation with interesting problems to learn more about them. He may then perhaps appreciate that they are more difficult and less cut and dried than he might have supposed. Professor George could write a book which explains simply and carefully just what is at stake. It is badly needed.

P. C. WASON

Symptoms of Psychopathology: A Handbook. Edited by C. G. Costello. New York: Wiley, 1970. Pp. ix + 679. £7·50.

The first edition of Jaspers' *Psychopathology* appeared over 50 years ago, and yet it still sets a standard by which other books on the subject must be judged. When Jaspers first wrote, psychology had rather little to contribute to the study of abnormal psychic phenomena; and even in 1959, in the seventh edition, he could still write: 'theoretically, psychology is as necessary for the psychopathologist as physiology for the pathologist, but in fact we find many instances to the contrary. This is due to the fact that psychopathologists are concerned with material of which the normal counterpart has not yet been studied by the psychologist'. (Translation of Hoenig and Hamilton, 1962.)

Psychopathology has a variety of meanings. But whatever the shade of meaning chosen by a particular author, the foundation of any work must surely be the careful description of morbid psychic phenomena. Only if this is sound can the author move on to a worthwhile discussion; and this is true whether he chooses to concentrate mainly on explanatory theories, on methods of assessment, or on treatment. Jaspers' book is still pre-eminent precisely because its opening pages contain one of the most clear and thorough descriptions of morbid psychic phenomena which has been written.

These, then, are some of the problems which a new book on psychopathology has to meet. At first sight, Costello's handbook appears all set to disprove Jaspers' gloomy view about the value of the contribution of psychology to psychopathology. A work of nearly 700 pages, written by a team of 33 experts and with an author index running to over 2500 entries, it promises exciting reading. It comes very near to providing it, but it does not quite succeed. In part this is because Jaspers is, unfortunately, still correct: despite all the advances in normal and abnormal psychology there are certain classes of abnormal psychic phenomena which are not illuminated in any obvious way by the findings of psychological experiments. Even the most scholarly reviewer cannot remedy this and there are places where experiments are described which can be linked with psychopathological states in man only by rather loose analogy.

Unfortunately, there are some other places where the contributors fail to avoid the second kind of fault. They launch into a discussion of theories and of supposedly relevant psychological findings after only the most perfunctory account of the morbid phenomena which they seek to explain. When this happens it is inevitable that their account of theories becomes undisciplined and the discussion of assessment and treatment appears less than wholly realistic. Fortunately, more contributors avoid these problems than fall foul of them. Among the best contributors are those of Gwynne Jones on stuttering, Inglis on memory disorder, Kales and Berger on sleep, Barclay Martin and Sroufe on anxiety, and Meyer and Crisp on phobias. There are many others.

This book must, by any reasonable standards, be judged a success. Anyone concerned with the contribution which psychology can make to the problems of psychopathology will find much to interest him—and, despite its high price, will probably wish to see it on his bookshelf.

MICHAEL GELDER

Fears and Phobias. By I. M. Marks. London: Heinemann, 1969. Pp. 302. £2.25.

Wolpe's text *Psychotherapy by Reciprocal Inhibition* appeared in 1958, and provided a major stimulus for the nascent behaviour therapy movement. In the decade since then several major texts and collections of papers have appeared, and the journal *Behaviour Research and Therapy* is now solidly established. British behaviour therapists have tended to concentrate on the desensitization of phobias, and aversion from sexual deviation; the Americans have been more concerned with operant techniques for chronic psychotics and children, analogue studies of fears in non-patient groups, and the exploration of alternatives to desensitization, such as modelling and implosion. The original, vigorous espousal of behaviour therapy by some leading experimental abnormal psychologists was seen by them as complementary to an earlier demolition of the claims of psychotherapy to therapeutic efficacy and as heralding the development of a truly scientifically based discipline of behavioural change, independent of traditional psychiatry and of psychiatrists. In practice, the field has developed rather differently, particularly in Britain, as is shown by Isaac Marks' important book. Marks is a psychiatrist, as are several of his Maudsley colleagues who have carried out a sustained series of controlled trials of desensitization in the treatment of phobias. They have shown that clinical psychiatry is flexible enough to encompass new developments, and that fraternal strife is not a necessary element for the advancement of knowledge. However, it cannot be said that Marks' sophistication in experimental psychology matches completely the depth of his knowledge of clinical psychiatry. For instance, he makes little or no mention of Neal Miller's important analyses of approach–avoidance conflict, of Mandler's work on individual differences in autonomic self-perception, or of recent experiment and theory on the incubation of behavioural experiences by Eysenck, Wasserman and others, all highly relevant to the explanation of phobic behaviour. An investigation of individual differences in the frequency and intensity of incubation is likely to be fruitful, and Byrne's studies of repression–sensitization are very germane.

The range of psychological work to which Marks does refer provides ample evidence of the need for an increased emphasis on psychology in the training both of medical students and of psychiatrists. His book is an excellent example of the 'Maudsley style'; concise, measured and thorough, and is a most valuable compendium of the present state of psychiatric knowledge concerning the phobias. The emphasis is on evidence rather than speculation—perhaps too much so, and the book would have gained greatly by an integrative chapter on future research needs and directions.

It is salutary to be reminded that only 3 per cent of Maudsley out-patients present with a primary complaint of phobic disorder, and over half of these are agoraphobics for whom desensitization is relatively ineffective. Very recent work by Marks, described by him at the 1969 Dublin Conference on Behaviour Modification, strongly suggests that implosion, or flooding, may be much more effective for agoraphobics than desensitization. The theoretical rationale for flooding is not immediately apparent, although Solomon's findings, on the necessity for the *forced* extinction of well-learned avoidance in dogs, may be relevant.

With the exception of animal phobias—again it is valuable to be reminded that they number only 3 per cent of all phobic referrals—the age of onset of phobias is *post*-adolescent—hence prophylaxis in childhood is unlikely to be effective. The social setting of the origin of phobias is likely to be worth looking at—as is the association of phobic behaviour with unfavourable self-perceptions, and the consequent pyramiding, partly through cognitive rehearsal, of the complex of stimuli to which the patient becomes phobic. Marks seems unaware of contemporary research in cognitive social psychology, and restricts his discussion of such 'symbolic' aspects to psychoanalytic contributions.

As well as controlled studies of desensitization, the Maudsley group, particularly Lader and Mathews, have studied the physiological predictors and concomitants of therapeutic change, and this work is clearly and pertinently described. Phobic patients with circumscribed single symptoms acquire the conditioned eyeblink response more rapidly and extinguish it more slowly than agoraphobics, whose symptoms are multiple, a finding which is rather embarrassing to a learning explanation of the aetiology of phobias. Another unresolved problem is the possibility, raised by Mathews at the Dublin Conference, that the attainment of imagery is easier in relaxation, raising serious doubts about the interpretation of studies comparing various combinations of relaxation, graded hierarchies, etc. Eventually the therapist, or experimenter, will require external control over the frequency and intensity of the stimuli which now are presented in fantasy.

Marks argues for the importance, both of the patient–therapist relationship, even in the straightforward desensitization of circumscribed phobias, and the necessity of dealing with a patient's interpersonal difficulties, where these exist. He states 'desensitization and insight psychotherapy can thus be complementary procedures in certain cases and are not mutually exclusive'. However, 'insight' has no standing as an operational term and is as outdated a concept as the Gestaltist's 'ah-hah' experience. Current research on attitude change (see Goldstein's excellent text *Psychotherapy and the Psychology of Behaviour Change*) is *at least* as applicable to the manipulation of the therapy relationship and of the self- and other-perceptions of the patient, as are the animal studies of approach–avoidance learning to desensitization and aversion. The next 10 years should see the systematic application of experimental social psychology to therapy, and there is an urgent need to include this area of psychology in the training, both of psychiatrists and of clinical psychologists. Unfortunately, many of the former seem disposed to climb uncritically on the ethology bandwaggon, relevant though much of animal social behaviour is, to the consequent neglect of research findings on human interpersonal behaviour.

Marks has made a most useful contribution to the growing literature on behaviour modification, and his book should be read by all concerned with the therapy of neurotic disorders and research into their causes.

M. P. FELDMAN

A Handbook of Contemporary Soviet Psychology. Edited by Michael Cole and Irving Maltzman. New York: Basic Books. 1969, Pp. xviii + 887. $25.00.

To the conscious reviewer, this handbook seems like Shakespeare's star 'whose worth unknown although his height be taken'. It is easy to ascertain the price of the book (high), its weight (considerable) and its length (tremendous); it is rather more difficult to digest and discuss the 30 separate chapters, which are grouped into four sections: Developmental Psychology, Abnormal and Social Psychology, General Experimental Psychology, and Higher Nervous Activity. What is quite obvious, even after a cursory glance, is that anyone concerned with any of the areas mentioned, and who wishes to gain some understanding of Russian work and Russian thinking, would have little alternative but to get hold of this book and read it with the utmost care and concentration. The editors have assembled a team which could hardly be improved; all are experts in their various fields, and many are familiar to us through their translations or through personal contacts. In addition, and this is an important point, the book has been edited with the utmost care. The editors have contributed an invaluable 38-page introduction which sets the scene and destroys many false preconceptions; in addition they have prefaced each chapter with a brief but informative note about the author, his standing, his academic provenance, and anything else that might help the reader to understand the chapter. Quite often they have added footnotes to draw attention to similarities or dissimilarities noticed between Russian and Western studies, or drawing the attention of the reader to other parts of the book where relevant issues are being discussed. In these days, when editing often simply means putting your name on the title page and sending off the manuscripts received to the printer, these points are worthy of being emphasized; they make what would otherwise be a good book into an exceptional one.

Among the misconceptions of Soviet psychology which many people hold is that of its exclusively Pavlovian orientation. This is what the editors have to say: 'A common misinterpretation of Soviet psychology is that it employs, almost exclusively, classical Pavlovian conditioning methods, and that all phenomena are forced into a common theoretical framework which views behaviour as the sum of atomistic, elementaristic reflexes. Psychology in the U.S.S.R. was never like this, was never restricted to Pavlovian conditioning, and certainly is not today. Pavlov's theories are influential, but these—and the empirical data upon which they are based—are far more complex than suggested by the typical American textbook treatment. An examination of current Soviet research indicates that relatively little classical conditioning is being conducted, much less than in the United States. It has always been done primarily by physiologists, those investigators concerned with "higher nervous activity". But behavioural conditioning studies are on the decline even among these physiologists. Since they are trained as physiologists and their principal interest is the function of the brain and its relation to behaviour, they are employing modern electrophysiological techniques for investigating brain activity. They are no longer restricted by the lack of adequate technology, as in Pavlov's time, to behavioural studies which could provide only a relatively gross basis for inferring brain activity.'

The chapters have been translated very well indeed; again, this is the exception rather than the rule. Many people have been put off Russian work for life by reading some of the atrocious translations that have been so common. Within the limits of the possible, these translations are readable, written in decent English, and psychologically meaningful in so far as any translation can achieve this difficult aim. Authors, editors and translators alike are to be congratulated on what constitutes a tremendous achievement the difficulties of which can only begin to be realized by someone who has had to edit a large cooperative book, and who has dealt with Russian material in the past.

The reader will find, in spite of all the good features of the book here emphasized, that the going is pretty hard. The editors explain in their introduction why this should be so; there are general differences in style of working and in mode of presentation which distinguish Russian from Western work, and even bearing these differences in mind does not make the translation easy. A word or a concept might have quite different meanings to a Russian and an American psychologist, and it is not always possible to reconcile these meanings by a suitable translation. The fact that the chapters are rather short does not help; if the authors had had more space they might have been able to explain more clearly. Possibly the editors tried to pack in too much; it might have been better to have omitted some of the more esoteric and applied topics.

Readers might be tempted to try out some of the techniques pioneered by Russian workers, and described here. They may find considerable difficulty in doing so. In chapter 18 is mentioned the photochemical conditioned reflex; this is used as a measure of the Pavlovian concept of 'mobility' in Teplov's system of personality. One of my students spent several years trying to obtain this phenomenon, and finally thought that he had probably succeeded; however, the difficulties involved and the artifacts inevitably associated with the phenomenon were such that he could not be certain. In any case, to use this as a personality measure was completely impossible. The authors of the chapter mention in passing that weeks or months of investigation may be necessary with each subject before worthwhile results can be obtained; this may serve as a warning to investigators whose subjects may run out of patience.

All in all then, this is an excellent but difficult book; if any volume can give us an indication of what Russian psychology is like at the moment, this must be it. But the 'if' remains; the task may be an impossible one. Even so, this book certainly takes us part of the way, and, as the Russians say, 'Better half a loaf than no bread!'

H. J. EYSENCK

Honest Sex. By Rustum and Della Roy. London: Allen & Unwin, 1969. Pp. 209. £1.75.

This is not an easy book to review objectively, largely because it confronts the reader with ideas that provoke intensely personal responses, but therein lies, perhaps, the book's greatest merit. In an introductory warning and plea, the authors exhort the reader to persist in his reading with an open mind and as much good-will as he can muster, even if the ideas are found to be offensive or even preposterous. Readers are requested also, not to quote from the book without reading most of it; a wise precaution. Having read the whole book, few will feel as outraged as they might have felt had they merely dipped into random sections.

The authors, a married couple, are both science teachers at the State College, Pennsylvania, and are active Christian laymen. Rustum Roy helped to found the Sycamore Community, an experimental Christian group, in which took place the exploration of the ideas formulated in this work.

The message is spelt out clearly many times. In an age of increasing sexual affluence, American society urgently needs new guidelines for its sexual morality. 'In other words, the traditional sources of morality and moral judgement have failed, in this generation, to come up with responsible do's or don't's for the citizen of secular society. Current thinking about sex standards is a chaos or a kaleidoscope or a morass.' Major responsibility is placed, by the Roys, at the door of the Church for having neglected sex almost entirely. A strong case is mounted to support the view that there is no literal Biblical basis for society's mores on such questions as monogamy. Following from this, the authors set out to formulate a new Christian sex ethic which is based on a number of explicitly stated views. Sex is seen as 'a continuum in depth and intensity from the gentle tingle of pleasure in an appropriately appreciative glance to the ecstasies of prolonged and diversified coitus' and it is regarded as both a source of pleasure and a language for communicating the deepest love between man and woman. The emphasis on the depth of relationship is fundamental and it follows that any particular level of sexual intimacy is justified only by the existence of appropriately deep and continuing relationships. Thus the depth of relationship appropriate for coitus is one that resembles one that is customarily needed for marriage. The recurring underlying principle is bound up with the somewhat idealistic concept of the supremacy of Agape—unselfish love for the other.

In the light of these views, the currently prevailing moral attitudes towards a wide range of human sexual feelings and expressions are examined. Traditional monogamy and monogyny are dealt with at some length in what is perhaps the most controversial section of the book. The authors expound the view that 'Men and women were not created to belong exclusively to each other in a marriage contract. They belong only to God and to all of Man'. Sex and marriage, therefore, are not necessarily exclusively linked, and 'once the conclusion is reached that the Christian ideal of unselfish Agape, concern for all, does not restrict all sexual expression to marriage, many obvious yet radical changes become possible'. Such changes are discussed in the context of various patterns of sexual relations, both inside and outside marriage.

The chapter on single women is particularly significant and is appropriately thought-provoking. Other chapters deal with the questions of premarital sex, divorce, birth control, homosexuality and 'solo sex'.

This inevitably controversial book is not, nor was it intended to be, the definitive work on a new Christian sex morality. It is, however, sincere, provocative and penetrating, without being righteous or unrealistically idealistic. The authors indicate that they owe much to, amongst others, Bishop J. A. T. Robinson and *Honest Sex* fairly deserves to have some of the impact that was made by *Honest to God*. There is a usefully annotated bibliography.

W. L. PARRY-JONES

Personality Tests and Reviews. Edited by O. K. Buros. Highland Park, N.J.: Gryphon Press, 1970. $35.00.

This book is an invaluable addition to the many Mental Measurements Yearbooks and other texts which Buros has published since the early 1930s; it contains detailed information and reviews by experts of a vast number of personality tests, many of which even an expert might be forgiven for not knowing. Also given with each test is a bibliography of articles in which it has been used; where I have been able to check these bibliographies they have invariably been accurate and reasonably complete. (By saying 'reasonably' I am indicating that some rather esoteric journals published in India, Japan, Poland, etc. have occasionally been omitted; this is inevitable and does not constitute a criticism of the book. Absolute completeness is unobtainable, as even the most obsessional of us must have realized.) The reviews are usually very much to the point, contain valuable criticisms and enable the reader to form a reasonable judgment of the strength and weakness of a given test.

It is difficult to say any more about a book of this kind without going into irrelevant and niggling detail. It is obvious that every psychological and psychiatric library must have a copy, and so, one would imagine, must everyone else actually working with personality tests. One of the values of the Buros series has always been that the critical comments of the reviewers would lead users into abandoning the less soundly based tests, and would encourage the producers of tests to mend their ways. Whether such a desirable outcome will in fact result is difficult to say; it is quite clear from the reviews that the majority of tests are still well below the level of competence in construction and validity which would mark them out as mediocre; far too many of even widely used tests are clearly bad and should be abandoned. This book is an important contribution to psychology; let us hope psychologists will make optimal use of it.

H. J. EYSENCK

Discussions on Child Development. Four volumes. Edited by J. M. Tanner and Bärbel Inhelder. London: Tavistock Publications. Pp. 240, 271, 223, 186. £1.50. per volume; paper, 60 p.

These four volumes are a reissue of the proceedings of the WHO Study Group on the Psychobiological Development of the Child, originally published between 1956 and 1960. A wide range of disciplines, from psychoanalysis through psychology and ethology to biology, psychiatry, electrophysiology and general systems theory was represented in the permanent members of the group and in their guest speakers. Particularly at this point in time, the value of the series does not come from the content of any of the presentations (although it is interesting to note that the 'softer' a speaker's background, the less dated his comments seem), but rather from the form the discussions took. Two quotations from the chairman's introduction to the first session give the tone of the whole enterprise: 'We tried to bring the discussions that normally take place in the corridor outside the scientific meetings into the meetings themselves. . . . We try to encourage the participants in our meetings to speak freely, not to hesitate to ask a foolish question, because how can they be sure it won't evoke wisdom in someone else? We are not here in this kind of group to solve a point, to defend a view, but rather to examine our own blind spots, or at least to give others a chance to point them out to us.' Thus there is comparatively little formal argument but a continual give-and-take between experts in different fields familiar with each other's concepts and methods and so with the different aspects of the processes of growth and development. As a result, as well as the exposition of fundamentals inseparable from genuine education, there is much incidental insight into the chaotic world of intractable men, machines and animals which everyone knows underlies the placid surface of published results but which is never publicly acknowledged, and into the formative influences on the development of the study of psychobiological development. In terms of current relevance perhaps Volume IV, with Piaget's statement of general problems and the comments on it of Lorenz, Bowlby, Mead, Grey Walter, Tanner, Zazzo and von Bertalanffy is the major contribution; but this is not the only criterion which should be applied. The breadth and informality of the discussions, the almost visible evolution of understanding in first-class intellects, determined to cooperate with each other in approaching problems held in common, and the enthusiasm with which they go about it, make this series the most entertaining as well as stimulating way of extending one's own knowledge of the whole area. One is left envying the participants their good fortune, a feeling mitigated by the consolations of this account of their interactions.

J. A. TAYLOR

A Textbook of Psychosexual Disorders. Second edition. By Clifford Allen. London: Oxford University Press, 1969. Pp. viii + 478. £3.50.

The first edition of this textbook appeared in 1962, although this grew out of an even earlier work first published in 1940. The author was for many years consultant psychiatrist to the Dreadnought Seaman's Hospital as well as in private practice. He has drawn on his extensive clinical experience and quotes many interesting case histories; he is clearly a humane person and has read widely.

This second edition is 70 pages longer than the first, but is divided into the same eight sections with the same titles: (1) The Sexual Instinct and its Development; (2) Disorders of Sexual Expression; (3) Disorders of the Instinctual Object; (4) Disorders of the Sexual Stimulus; (5) Disorders of the Instinctual Strength; (6) Prostitution; (7) The Prevention, Treatment and Prognosis of the Psychosexual Disorders; (8) The Medico-legal Aspects of Psychosexual Disorders.

In the preface the author rather disingenuously states: 'No attempt has been made to pile up an unwieldy mass of facts to confirm what seems already proved. Where proof and descriptions seem adequate, the book has not been expanded; where new knowledge has been revealed, or wider descriptions made possible, no hesitation has prevented expansion.' He also notes that this is the only textbook of psychosexual disorders in any language. Unfortunately it is also the worst.

It lacks scholarly perspective and scientific objectivity. Fundamental problems are either ignored altogether or simply glossed over, and clinical anecdote is too readily accepted as the final arbiter. The subject is not historically evaluated—how should the work of the pioneer sexologists, e.g. Krafft-Ebing, Havelock Ellis, Hirschfeld, be seen today? There is nothing on the fundamental differences between the sexes, concepts of masculinity–femininity, and the word 'gender' never appears at all.

Problems of definition and terminology are arbitrarily and superficially dealt with. Terms such as perversion, inversion, aberration and paraphilia are loosely used and seldom properly defined, especially with any regard to socio-cultural factors. In fact, modern sociological contributions, e.g. Schofield on homosexuality, are ignored. Whilst some attempt is made (p. 77) to define some terms, many others are left out; a comprehensive glossary with derivations where known would be much more useful. That 'onanism' should still be used is surprising in a work of this kind and the fact that 'voluntary treatment', 'certified' patients and suchlike still appear bears witness to inadequate revision.

That some literary references are given is all to the good—it helps to leaven the clinical aspects, widens the scope and makes more interesting reading—but these are so few. Is there no place for Swinburne, T. E. Lawrence and to mention only two first-rate scholarly surveys, Jeannette Foster's work on *Sex Variant Women in Literature* and Mario Praz on the *Romantic Agony*?

The chapter on endocrinology is hopelessly out of date and has nothing on the recent most interesting and important advances in neuro-endocrinology.

That homosexuality should receive the most extensive coverage is understandable due to the fact that it is so common and a great deal of research has been carried out. However, the author almost totally ignores female homosexuality (and indeed throughout the book shows little interest in or experience of psychosexual problems in women). Nor does he carefully distinguish the paedophiliac. That he has written a separate monograph on homosexuality is not relevant, as it is equally deficient.

General assessment receives a separate section towards the end of the book (it would have been more logical to have it earlier on) but so preoccupied is he with psychodynamics that the most obvious and most important basis for this, viz. a comprehensive psychiatric history and examination of mental state, is barely mentioned. There is no adequate discussion of the place of the clinical psychologist or psychometrics in either the initial assessment, as an aid to therapy, measurement of change or in research.

Under coprolalia there is no mention of Gilles de la Tourette's syndrome. In discussing sadomasochism, the term compulsion is introduced but this really confuses the issue; he misses the essential point of the inner struggle and overcoming of resistance as well as the fact that patients with true compulsions do not enjoy them. The opportunity is not taken of exploring the relationships between sexual deviations and personality disorders, particularly psychopathy. Flagellation does not receive the attention it deserves and the section on masochism and suicide is very muddled.

Out of the many unsubstantiated statements the one on the effects of exhibitionism on children could be taken as an example (p. 187): 'it can frighten children sexually and have a permanent effect upon their later life'. There is no evidence for this. Other examples could be given.

The definition of transvestism is poor both in substance and in logic. He seems unaware that women as well as men seek sexual satisfaction out of their peer group and frequently 'beneath' them (what might be called the Lady Chatterley complex). This need not necessarily have a very subtle psychodynamic meaning but may merely express a more general inadequacy or rebellion against authority and the established order. Throughout the author relies far too heavily on Freudian dynamics and explains practically every condition by involving the mother.

This can sometimes lead to neglect of the most common associated cause, e.g. loss of sexual desire in a depressive illness or the fact that certain newer drugs can cause impotence. Psychoanalytical formulations are slavishly followed even when there is no evidence to support them or when they have been greatly modified by analysts themselves, e.g. (p. 355) a woman with sexual hyperversion is 'frequently to some extent homosexual, and thus behaves in an active rather than a passive way'. The common association with hypomania is not noted and he seems never to have heard of de Clérambault's syndrome. But perhaps an author can be forgiven for missing the rareties, but what of jealousy—surely a very common complaint—which is not mentioned at all?

The chapters on prognosis (mainly on homosexuality) and treatment are both very poor. He constantly advocates *private* psychotherapy as the treatment of choice, has little faith in short-term methods carried out at out-patient clinics or for the use of drugs. There is too much special pleading here and very little to back it up with; for instance, he produces no statistics or follow-up data on any series of patients treated by himself. Not surprisingly, he has little to say on behaviour therapy and some of this is wrong; e.g. Wolpe was by no means the first to develop progressive muscular relaxation.

The references are grouped together at the end of chapters—they are frequently out of date, arbitrarily presented (e.g. sometimes a paper is given a title, sometimes not) and poorly checked. The index is quite inadequate, some conditions are mentioned in the text and are not indexed at all, e.g. bondage, pygmalionism, whilst others such as arson do not appear at all. There should be a separate name index.

It is a great pity that a second edition was brought out without a very radical revision. In its present form it cannot be recommended as a reliable source of information on the psychosexual disorders.

F. E. KENYON

Man and Woman: The Basis of Their Relationship. By K. H. Wrage. London: Collins, 1969. Pp. 259. £3.50.

The author is director of the Bureau of Social Services and Medicine of a German Protestant Church. The book was originally published in German in 1966. The present translation is by S. S. B. Gilder, editor of the *World Medical Journal*. It is a handsome volume, beautifully produced and well illustrated by G. Kapitzke. It is intended as a comprehensive manual of sex education.

In spite of efforts at scientific objectivity and liberal sentiments the 'Protestant ethic' is well in evidence. It is not always up to date or explicit enough, e.g. nothing on the Abortion Act 1967, and very little indeed on any pathology, either psycho or somatic. However, most of the information presented is concise, accurate, simply stated and whenever possible amply illustrated. There is a glossary of technical terms, a short reading list and an adequate index.

It would be a useful book for self-education and for recommending to selected patients, as well as for marriage guidance counsellors and others concerned with education and guidance, but rather too elementary for other professional workers.

F. E. KENYON

Psychosomatic Aspects of Gynecological Disorders: Seven Psychoanalytic Case Studies. By A. O. Ludwig, B. J. Murawski and S. H. Sturgis. Massachusetts: Harvard University Press, 1969. Pp. 119. £2.40.

The authors of this monograph are respectively psychiatrist, psychologist and gynaecologist at the Harvard Medical School. They set out to study the influence of emotional stress and personality structure in gynaecological disorders by intensive study of seven patients using projective psychological tests, psychoanalysis and measurement of adrenal steroid excretion (17-ketosteroids and 17-hydroxycorticosteroids).

Patients were studied for periods varying from 3 to 7 years and they completed anything from 186 to 950 hours of psychoanalysis. A total of 1563 biochemical tests were done on 24-hour specimens of urine collected as far as possible twice weekly. During the course of the investigation three patients had full-term pregnancies. Projective tests used were the TAT and Rorschach with the aim of testing both before and after treatment. No normal controls were used as none could be found willing to supply the necessary urine samples.

The bulk of the book consists of extended case reports. Case 1 was 24 and suffered from amenorrhoea and asthma; the latter got worse during the period of study but it is not made clear what physical therapy was also given. Case 2, aged 32, had a 14-year history of amenorrhoea but was subsequently shown to have polycystic ovaries. Case 3 was 37 with severe premenstrual tension, menorrhagia, iron-deficiency anaemia, and chronic mastitis. She also had difficulties with interpersonal relationships, especially with her husband and children.

Case 5 was 25 and had hysterical and obsessional personality traits as well as suffering from depression and frigidity. Case 6 was 26 and had two main complaints, loss of libido and vaginal discomfort. Case 7 was 25 and diagnosed puerperal psychosis; at times she was deluded, hyperactive, euphoric and unmanageable. Treatment with drugs (chlorpromazine) was necessary.

The overall conclusions are discussed in a final summarizing section. Each patient had a distinctive pattern of steroid excretion and some relationship was established between high excretion rates and stress, particularly separation or threat of separation from the analyst. In general, steroid excretion levelled off in the later phases of the analysis.

Appendix 1 consists of notes on the psychological test results, and Appendix 2 further details of Rorschach responses, but not the complete protocols. Test data were incomplete in several instances, e.g. none in Case 3 to preserve confidentiality, and two others did not have post-treatment testing. There are few references and no index.

It is very easy to criticize this type of study, but very difficult in practice to overcome the more obvious shortcomings. The methodological problems are formidable, and perhaps the main value of this study lies in the highlighting of these rather than in the deductions drawn from them.

F. E. KENYON

Adolescent Psychiatry. Edited by S. J. Shamsie. Quebec: Schering Corporation, 1969. Pp. 84. No price given.

This slim volume is the outcome of a conference in Montreal at which several psychiatrists presented papers about various aspects of adolescence as a period of growth and about adolescents who presented particular clinical and social problems. The contributors were all agreed that many psychiatrists had avoided adolescent psychiatry, and, for this reason, principally they were intent upon creating a sub-speciality that would concentrate upon adolescent disorders. However, from the range of papers in the book it was clear that psychiatrists have one contribution to make, and that physicians, psychologists, educators and social workers have others. The book might have had more merit if people from the 'paramedical' fields had themselves presented papers on their traditional areas of concern, perhaps to the surprise of those who thought that such people could be trained to handle adolescents. In fact, by their own admission it seems that psychiatrists need the special training for contacting disturbed adolescents, for bringing them into treatment, and for encouraging them to stay in treatment.

In general, this book gives a good, clear statement of many of the problems presented by adolescents, and some contributors were more compelling with their material and their style of writing than others. I am still reflecting on the comment that 'it is hard to conceive of a way in which an emotionally disturbed adolescent can show behaviour without his violating some municipal, provincial or federal law'—and I wonder why adolescents do not have the range of neurotic disorders and symptoms through which to express their conflicts. I was disappointed not to find a paper on drug addiction in the book because of its frequency among adolescents in North America, and the one paragraph on the prescription of drugs and the abuse of the same could well have been extended without seeming to defer to the sponsors of the conference—a drug company. Pupil and student demonstrations and protests might also have been mentioned. Finally, the concluding squabble between those who enjoy adolescent psychiatry and those who don't would suggest that the administrators might allow freedom of choice to both sides to follow their preference so long as adolescent patients are not neglected in the process—the case for a sub-speciality of psychiatry was 'not proven'.

A. J. W. TAYLOR

Drugs from A to Z: A Dictionary. By Richard R. Lingeman. London: Allen Lane The Penguin Press, 1970. Pp. 263. £2.50.

Public ignorance and the blanket condemnation of illicit drug use persist. Perhaps the stereotype of the wild-eyed 'fiend' has been replaced by that of the sullen, unkempt, resentful youth, but in Britain, at any rate, the hysterical reception, even among educated and presumably responsible public figures, of the Wootton Committee's Report on marihuana use testifies to the apparently inexhaustible reserves of moral indignation among those social strata with the power and influence to make their morality effectively everybody's morality.

Given such an agitated climate of opinion, a dictionary on drugs is perhaps an obvious sedative. Lingeman's volume is concerned exclusively with drugs of abuse and their substitutes and antagonists; it is a true dictionary in the sense that the entries cover derivation, synonyms and, in many cases, illustrations of literary usage as well as the cross-referenced definitions. The latter are relatively lengthy for items of central interest: 'marihuana' gets nine pages, 'heroin' eight. It is possible to find out what an 'ice cream habit' is, or trace 'meserole' to one Mezz Mezzrow of Harlem's preference for fat reefers. The coverage of substances is wide (although omissions include Antabuse and Temposil) and the background information on modes and consequences of use is judiciously selected.

Although revisions have been made for the English edition (including British brand names and equivalents) the ethnocentricity of some entries is exacerbated by their inconsequentiality. That 'scissors' has been slang at Shimer College, Mt. Carroll, Ill., for marihuana, or that 'stuff back' meant 'to eat a lot' in 1946, casts a dimmer light on the user's world than the author's preface suggests. The inclusion of standard dosages in a book intended primarily for a lay audience is questionable and the compiler's disclaimer concerning medically supervised drug taking, while no doubt sincere, remains a little lame. Although a serious intention can be discerned underlying Lingeman's effort, its potential influence on opinion in Britain will be severely curtailed by the content of various entries and quotations, which to some serious judges will appear largely irrelevant and to others unacceptably flippant.

KEN LEVINE

CHANGE-OVER TO SI UNITS

From January 1972 the journals of the British Psychological Society will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Further information about SI Units is contained in the revised *Suggestions to Authors* pamphlet issued by the Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

Authors who refer to physical measurements in their papers should now use SI Units except where it would be absurd to do so. Conversion tables will be published during the change-over period but in *Br. J. soc. clin. Psychol.* these will be restricted to units of length as others rarely occur. When they do, conversion factors will be stated in footnotes. Common units of time (e.g. hour, year) will persist.

1 metre (m) = 39.37 inches
1 metre (m) = 3.28 feet
0.0254m = 1 inch
0.3048m = 1 foot

Br. J. soc. clin. Psychol. (1971), 10, pp. 301–305
Printed in Great Britain

The Effects of Subject-Generated and Normative Constructs Upon the Formation of Impressions

By JESSE G. DELIA
University of Illinois

ANDREW H. GONYEA
Clark University

AND WALTER H. CROCKETT
University of Kansas

This study examined the validity of G. Kelly's Individuality Corollary by assessing the role of individual personal constructs in the formation of impressions. All subjects formed impressions from six traits, three positive and three negative, attributed to an unknown person. All of the traits presented to one group were normatively determined. Three other groups received six traits partly or entirely generated by the subjects themselves in descriptions of other persons. These groups were characterized as own-positive, own-negative and all own. As hypothesized, subjects produced more differentiated impressions in the conditions in which they received as stimuli the self-generated traits. Evaluations of the unknown other moved in the direction of the evaluative component of the self-generated traits. In addition, the most highly organized impressions were produced in the normative and all-own conditions. These results were interpreted as supporting both the Individuality Corollary and a conflict model of impression organization.

George Kelly's theoretical system (1955) is especially appropriate for analysing the cognitive processes that influence one person's impression of another's personality. Kelly proposed that man employs a set of personal constructs by which he construes, among other things, the character of other people. In Kelly's system, an individual's personal constructs are conceived as dimensions to be employed in interpreting his experience. An event makes sense only within the framework of the perceiver's system of constructs.

Kelly's conceptions have been widely adopted in theoretical accounts of how individuals process information in forming impressions of others (cf. Bannister & Mair, 1969; Bieri *et al.*, 1966; Crockett, 1965). Studies of impression formation, however, have not examined the relationship between the specific content of a subject's construct system and the use he makes of the information that is available to him about the other person. In his Individuality Corollary, Kelly emphasized that each individual develops a set of constructs that is unique in its content and organization. This implies that the extent to which a perceiver utilizes information about someone else should vary with the degree of correspondence between the nature of the information and the perceiver's own constructs.

The specific application that was made of the Individuality Corollary in the present study grew out of an interpretation by Crockett *et al.* (1970). These authors presented subjects with information that included bivalent, potentially contradictory qualities of another person. For some subjects the information was presented as abstract traits; for others it was presented in descriptions of concrete behaviours which represented the identical qualities. Subjects who received the abstract information, compared to those who received concrete behaviours, (*a*) formed more extensive impressions and (*b*) more often represented and reconciled in their impressions the contradictions in the information. It was concluded that subjects in the abstract group most often ordered the qualities directly to constructs in their own systems; in their impressions these subjects then had to deal directly with the contradictions and, if possible, to reconcile them. By contrast, subjects in the concrete group could deal with the contradictions at a different level, i.e. by interpreting the concrete behaviours so as to minimize the conflict among the constructs that were involved. With less conflict among constructs, subjects were not as strongly impelled to produce extensive, complexly organized impressions to account for the contradictions.

If abstract qualities are ordered more directly into a subject's construct system than concrete behaviours, it should also be true that statements which employ the subject's previously elicited constructs will be ordered into his construct system even more directly than normatively selected abstract qualities. Thus a subject who receives bivalent information expressed through his previously elicited constructs, compared to one who receives normatively selected abstract qualities, should form more highly differentiated, more complexly organized impressions. By the same token, when information of one valence is expressed through the subject's own constructs and that of the opposite valence through normatively selected ones, the impression should be dominated both in content and in valence by inferences from the subject's own constructs.

METHOD

Subjects

Subjects were 55 student volunteers enrolled in one of three freshman-level Fundamentals of Speech Communication classes at the University of Kansas.

Experimental procedure

The experimental paradigm was identical in all conditions. Subjects received an experimental communication consisting of a set of six statements, three positive and three negative, about a young man named John. The statements were allegedly made by six different people who knew John well. Positive and negative statements were presented in alternation. Each statement presented one quality, followed by a brief elaboration which specified its meaning. For example, the quality *conforming* was presented as follows: '*Conforming*: He behaves so as to seem to be in accordance with the expectations of others.' After reading over the statements, subjects were to imagine that a close friend desired to know as much as possible about John. Subjects wrote an impression expressing what they would tell their friend, then evaluated John on an 11-point scale, end-anchored by the terms 'very unfavourably' and 'very favourably'. Subsequently, subjects completed a brief post-impression questionnaire.

Experimental conditions

The study employed four experimental conditions as follows.

Normative-only. All subjects received the same six traits: *confident*, *conforming*, *idealistic*, *sarcastic*, *polite*, and *nosy*. Though evaluatively bivalent, these six traits are descriptively independent according to Peabody's (1967) factor analysis of trait dimensions. The evaluation of the six traits taken together proved to be approximately neutral on Anderson's (1968) likeability ratings of 555 trait names.

All-own. In an earlier session, subjects had described the kinds of men and women their own age they would like and dislike, and indicated which of the qualities in each description were most important. The stimulus material in the all-own condition consisted of three positive and three negative qualities chosen individually from each subject's own descriptions, using, if possible, only qualities he had labelled 'most important'. Each quality was followed by a brief elaboration, so that the form of the material corresponded exactly to that of the normative-only condition.

Own-positive. Three of the subject's own positive constructs, as determined above, were placed in alternation with the negative traits from the normative-only condition.

Own-negative. Three of the subject's own negative constructs were placed in alternation with the positive traits from the normative condition.

Dependent measures

Three major dependent variables were used in the study. The *degree of differentiation* of the impressions was defined as the total number of traits in the subject's description of John. This score was further divided into traits of positive and negative valence. Secondly, impressions were scored for *level of organization* by an extension of the developmental scoring scheme described by Kaplan & Crockett (1968). In this system,* impressions are scored at one of 15 levels depending upon the extent to which the contradiction is represented and reconciled. The third dependent variable was the subject's evaluation of John on an 11-point scale.

Summary of the data analysis

Of the 55 subjects, only 44 completed both experimental tasks. Four of these were randomly discarded to provide 10 subjects per cell.

Separate analyses of variance were conducted on the three dependent measures. For analysis of the degree of differentiation of the impressions, a 4 × 2 analysis was used, with repeated measures on the second factor. The two factors were type of information (normative-only, own-positive, own-negative, all-own) and within-valence differentiation (obtained by scoring the positive and negative constructs separately). One-way analyses were computed on the level of organization and evaluation scores. Since the all-own condition was irrelevant to the specific hypothesis being investigated, it was excluded from the analysis of evaluation scores.

RESULTS

As Table 1 shows, more differentiated impressions resulted from traits based on the subject's own constructs. The analysis of variance of differentiation scores revealed two significant effects: the main effect for type of information ($F = 3{\cdot}98$; d.f. = 3, 36; $P < 0{\cdot}025$) and the interaction between type of information and valence of the constructs ($F = 4{\cdot}52$; d.f. = 3, 36; $P < 0{\cdot}01$). Examination of the means for this interaction shows that subjects in the own-positive condition produced more

Table 1. *Mean differentiation, evaluation and level of organization of impressions in experimental groups*

	Experimental group			
Dependent variable	Normative-only	Own-positive	Own-negative	All-own
Positive differentiation	2·8	5·3	3·0	6·1
Negative differentiation	4·0	3·2	5·8	5·2
Total differentiation	7·6	9·3	8·8	11·6
Level of organization	10·5	7·8	8·8	11·4
Evaluation	5·5	7·1	4·6	6·5

---

* Duplicated copies of this manual are available from W. H. Crockett.

positive constructs than did those in the normative and own-negative conditions. Similarly, subjects in the own-negative condition produced more negative constructs than did those in the normative and own-positive conditions. In the all-own condition, positive differentiation approximated that in the own-positive condition, and negative differentiation approximated that in the own-negative condition. Clearly, then, the main effect for type of information was carried by the within-valence differentiation from the subject's own constructs. When the subjects received three normative traits of one valence, they produced an average of 3·25 constructs of that valence in their impressions. In contrast, when the three stimulus traits expressed subjects' previously elicited constructs, they produced an average of 5·58 constructs of that valence in their impressions.

The prediction that the stimulus person would be evaluated in the direction of the valence of the subject's own constructs was also supported. The overall analysis of variance on evaluation scores was significant at the 0·05 level ($F = 3{\cdot}36$; d.f. = 2, 27), and the comparison of the own-positive and own-negative groups was significant ($t = 2{\cdot}67$; d.f. = 18; $P < 0{\cdot}02$, two-tailed). The difference between the evaluation in the normative group and that in the own-positive group was nearly significant ($t = 1{\cdot}78$; d.f. = 18; $P < 0{\cdot}10$, two-tailed). However, the difference between the evaluation of the own-negative group and that of the normative-only group, while in the expected direction, was not significant ($t = 0{\cdot}93$; d.f. = 18).

The analysis of variance of level of organization scores also approached significance ($F = 2{\cdot}54$; d.f. = 3, 36; $P < 0{\cdot}075$) As Table 1 shows, this effect reflected the fact that the impressions in the all-own and normative conditions were organized at a more advanced level than those in the own-positive and own-negative conditions.

Responses to the post-impression questionnaire showed that in all conditions the stimulus person was believed to be a real person and that the subjects were familiar with someone like him. Two other results of interest from responses to the post-impression questionnaire are: (1) the stimulus person was seen as somewhat, though not significantly, more inconsistent in the normative and all-own conditions, and (2) when asked which traits were most important to their impression of John, subjects selected their own constructs at a frequency double that of the normative ones.

DISCUSSION

Consistent with Kelly's theory the results indicate that the basic processes involved in impression formation are the acts of interpretation and inference. Confronted with a body of information about another person, an individual initially 'makes sense' of that information by interpreting it within an existing system of personal constructs. Information which is construable directly within dimensions that he brings to the situation takes on great salience and is easily extended by inference to qualities that refer to closely related constructs. To the extent that the information which an individual receives is based directly on his own personal constructs, these cognitive acts are made easier. Since information based on one's own constructs directly implies the presence of related qualities, usually similar in valence to the elicited constructs, the evaluation of the stimulus person moves toward the pole of the own constructs.

The result for level of organization, reflecting higher organization in the all-own and normative conditions, suggests that potential contradiction in the information about another person may force the perceiver to a more complex level of understanding of that person. A number of previous studies using normatively selected qualities (Nidorf & Crockett, 1965; Rosenkrantz & Crockett, 1965; Meltzer *et al.*, 1966; Crockett *et al.*, 1970) have also suggested that one motivating force for organizing impressions is the need to subsume both positive and negative qualities in a unified impression. The present result suggests further that the level of organization attained is related to the *degree* of contradictoriness in the stimulus information. When information of only one valence was presented in the subjects' own constructs, many more qualities of that valence were inferred, thus dominating the impression, minimizing the contradiction, and removing any great need to provide a complex explanation of the ambivalence. In the all-own and normative conditions, on the other hand, the traits of opposite valence were nearly equal in weight and the contradictions were not easily minimized by bolstering one side of the information; hence, the achievement of a balanced, harmonious understanding of the other person required a complexly organized impression. A test of this interpretation will require research in which the degree of contradiction in the stimulus information is independently measured and systematically varied.

This research was supported in part by the National Institute of Mental Health, research grant number 1-R01MH-16563. Requests for reprints should be addressed to Jesse G. Delia, Department of Speech, University of Illinois, Urbana, Illinois.

REFERENCES

ANDERSON, N. H. (1968). Likeableness ratings of 555 personality-trait words. *J. Person. soc. Psychol.* **9**, 272–279.

BANNISTER, D. & MAIR, J. M. M. (1968). *The Evaluation of Personal Constructs*. New York: Academic Press.

BIERI, J. *et al.* (1966). *Clinical and Social Judgment: The Discrimination of Behavioral Information*. New York: Wiley.

CROCKETT, W. H. (1965). Cognitive complexity and impression formation. In B. Maher (ed.), *Advances in Experimental Personality Research*, vol. 2. New York: Academic Press.

CROCKETT, W. H., GONYEA, A. H. & DELIA, J. G. (1970). Cognitive complexity and the formation of impressions from abstract qualities or from concrete behaviors. *Proc. 78th Ann. Conv. Am. Psychol. Ass.*, 375–376.

KAPLAN, B. & CROCKETT, W. H. (1968). Developmental analysis of modes of resolution. In R. P. Abelson *et al.* (eds.), *The Cognitive Consistency Theories: A Sourcebook*. Chicago: Rand McNally.

KELLY, G. A. (1955). *The Psychology of Personal Constructs*. New York: Norton.

MELTZER, B., CROCKETT, W. H. & ROSENKRANTZ, P. S. (1966). Cognitive complexity, value congruity, and the integration of potentially incompatible information. *J. Person. soc. Psychol.* **4**, 338–343.

NIDORF, L. J. & CROCKETT, W. H. (1965). Cognitive complexity and the integration of conflicting information in written impressions. *J. soc. Psychol.* **66**, 165–169.

PEABODY, D. (1967). Trait inferences: evaluative and descriptive aspects. *J. Person. soc. Psychol. Monogr.* **7**, pt. 2. (Whole no. 644.)

ROSENKRANTZ, P. S. & CROCKETT, W. H. (1965). Some factors influencing the assimilation of disparate information in impression formation. *J. Person. soc. Psychol.* **2**, 397–402.

Manuscript received 8 June 1970

Revised manuscript received 7 January 1971

Br. J. soc. clin. Psychol. (1971), **10**, pp. 306–312
Printed in Great Britain

The Attribution of Behavioural Connotations to Shaded and White Figures by Caucasian Children

By GEOFFREY BROWN

University of Lancaster

and SUSAN P. JOHNSON

Danum Grammar School

Children between the ages of 3 and 11 years were asked to ascribe positive and negative behavioural characteristics to pictures of white and shaded figures. Whilst the younger children, 3 and 4 years, tended to do this in a random fashion, children between the ages of 5 and 8 years showed a marked increase in the attribution of negative statements to shaded figures and positive statements to white figures. From the age of 9 years there was a diminution in this behaviour. Exposure to coloured immigrant children in the school and play-group situations was associated with a decrease in this prejudicial behaviour.

There is much anecdotal material suggesting that the words 'black' and 'white' carry specific evaluative connotations. Isaacs (1963) noted that the use of black to symbolize evil and wickedness and white to symbolize goodness and purity pervades our culture.

Williams (1964) studied the connotative meanings of colour names by means of a form of semantic differential. For both Caucasian and Negro groups white was rated most positively on the evaluative dimension (good–bad) and black and grey were rated most negatively. He suggested that the regular association of the Negro with the term 'black' and the Caucasian with the term 'white' would tend to condition these characteristics to the respective races.

An investigation to see whether the connotative meanings of these two colours may be conditioned to terms with which they are regularly associated was conducted by Harbin & Williams (1966). Nonsense syllables and colour names were learned in a paired-associate learning schedule. Subsequent rating of the nonsense syllables on a semantic differential showed that those paired with black and white rated significantly differently on the evaluative dimension, and in the predicted directions.

A further study by Williams (1966) tested the hypothesis that the connotative meaning of a racial name would be more similar to the meaning of an associated colour than to colour names used to designate other groups. As predicted, American Caucasian students' semantic differential ratings indicated that triads of concepts linked by colour (e.g. white, white person, Caucasian) were significantly more close in connotative meaning than triads not so linked. A further study (Williams & Carter, 1967) indicates that this phenomenon is not restricted to the American culture.

Williams (1969) related his previous work to measures of racial attitude, estimated by four attitude-scaling techniques. Positive correlations were found between evaluative scores of the names 'black' and 'brown' and all attitude measures towards the Negro.

Renninger & Williams (1966) attempted to determine the age at which the concepts of 'white' as 'good' and 'black' as 'bad' were formed. They reported that Caucasian children were learning these relationships in their pre-school years, and only slightly in advance of the development of racial awareness.

Contradictory findings were presented by Williams & Roberson (1967). Using a modified form of the material devised by Renninger & Williams, they found evidence that racial awareness preceded the learning of colour meanings. This would suggest that colour-meaning factors reinforce the development of racial prejudice, whereas it had previously been postulated that 'the designation of racial groups by colour names may provide the child with a general evaluative frame of reference within which the more specific learning of prejudice may be easily incorporated' (Renninger & Williams, 1966, pp. 783).

An attempt to reverse the customary connotations of 'black' and 'white' was made by Williams & Edwards (1969), using operant conditioning techniques. Although no reversal was achieved the connotations were weakened. The study provides evidence in support of a link between colour concepts and racial attitudes, since children whose black–white concept attitude had been weakened showed a lower tendency to evaluate Negroes negatively and Caucasians positively.

Radke *et al.* (1950) gave experimental evidence of children's racial attitudes without linking these to colour connotation. They reported that children in all grades assigned more undesirable attitudes to Negroes and showed a strong preference for their own Caucasian group.

Stevenson & Stewart (1958) conducted a developmental study of racial awareness with children from 3 to 7 years. They reported that after the age of 3 years, at which responses were at random, there was a marked improvement in powers of discrimination, particularly amongst Caucasian children. This was accompanied by an increase in own-race rejection by Negroes.

The present study tests the hypothesis that, with increasing age, children more frequently attribute positive behavioural characteristics to non-shaded figures and negative characteristics to shaded figures.

METHOD

The investigation was conducted in the Midlands with 259 school children drawn from 37 primary schools, and 61 pre-school children from two play groups. The sample was made up of 160 of each sex and the ages ranged from 3·0 years to 10·11 years.

The sample was restricted to Caucasian children, as there is ample evidence that responses of Negroes are likely to display own-race rejection (Clark & Clark, 1947; Radke *et al.*, 1950). Investigation of such phenomena was outside the scope of this study.

The random sample of school children was drawn by selecting the second and third Caucasian boys, and the second and third Caucasian girls on each class register. The sample of pre-school children was obtained by using all the Caucasian children in the only available two play groups. The limitations of such sampling are acknowledged.

Subjects were tested individually, experimenter and subject facing each other across a small table or desk. Every effort was made to encourage the subject, though the words 'right' and 'wrong' were not used.

The apparatus consisted of two cards, on one of which were drawn identical outlines of two boys, and on the other were identical outlines of two girls. On each card one of the figures had the face, arms and legs shaded to a dark grey with soft pencil. The hair was coloured black on all figures.

The first card was presented to the subject with the following statement: 'Here is a picture of two boys. I am going to tell you some stories about them and I want you to point to the one you think the story is about.'

The following four statements were then read to the subject. After each one the subject indicated one of the figures on the card. (1) 'One of the boys has broken a window with a stone. Which boy do you think threw the stone?' (2) 'One of the boys always looks smart and tidy. Which do you think is the smart tidy boy?' (3) 'One of the boys is very naughty and pushed the other boy over. Which do you think is the naughty boy?' (4) 'One of the boys is very good and goes to the shop for his Mummy. Which do you think is the good boy?'

The second card was then presented with the following statement: 'Now I'm going to show you a picture of two girls. I am going to tell you some stories about them and I want you to point to the one you think the story is about.'

A similar procedure was then followed with the following statements: (1) 'One of the girls is a bully and hit the other one. Which do you think is the bad girl?' (2) 'One of the girls works very hard at school and the teacher likes her. Which do you think works hard?' (3) 'One of the girls spilled red ink on the floor and said that someone else had done it. Which do you think did this?' (4) 'One of the girls is very well-behaved. She never pushes to get on the bus but waits her turn. Which do you think is well-behaved?'

In both cases the first and third statements related to some form of negative behaviour and the second and fourth statements to some positive form of behaviour. The evaluative adjectives used are similar to those used by Radke *et al.* (1950).

The responses were scored to show the total number of positive and negative statements attributed to each figure.

RESULTS

The following scores were recorded: (1) the number of choices of the non-shaded figure when a positive statement was read; (2) the number of choices of the shaded figure when a positive statement was read; (3) the number of choices of the non-shaded figure when a negative statement was read; (4) the number of choices of the shaded figure when a negative statement was read.

The attribution of the specific non-shaded/positive and shaded/negative behaviour patterns will be referred to as a prejudiced choice, though this does not imply that the mechanisms in operation are necessarily identical with the mechanisms of racial prejudice.

For each of the eight age-groups Table 1 shows: (1) the number of times the non-shaded figure was chosen when a positive statement was read, expressed as a percentage of the total choices made to positive statements; (2) the number of times the shaded figure was chosen when a negative statement was read, expressed as a percentage of the total choices made to negative statements; and (3) the significance of differences between the choices of shaded and non-shaded figures with positive and negative statements, assessed by chi-square.

Table 1 shows that, with age, there was an increasing tendency to attribute positive statements to the non-shaded figures and negative statements to the shaded figures. It can be seen that the attribution of positive and negative statements to

Table 1. *Percentage of positive statements attributed to non-shaded figures and negative statements attributed to shaded figures in each age-group*

Age-group	Positive statements to non-shaded figures (%)	Negative statements to shaded figures (%)	Chi-square (with Yates' correction)	*P*
3–3·11	51·25	48·75	0·0	n.s.
4–4·11	51·25	50·63	0·05	n.s.
5–5·11	59·88	58·72	10·86	0·001
6–6·11	61·14	58·32	12·82	0·001
7–7·11	73·47	72·45	80·82	0·001
8–8·11	77·86	78·48	98·02	0·001
9–9·11	74·95	70·67	62·94	0·001
10–10·11	62·36	65·34	26·82	0·001

Possible sex differences within age-groups were also tested by chi-square; see Table 2.

the two figures was not significant in age-groups 3–3·11 and 4–4·11, but was highly significant ($P < 0{\cdot}001$) in all other age-groups.

As can be seen in Table 2, the sexes did not differ significantly in their choices, except in the 7–7·11 age-group, in which boys displayed a much greater number of prejudice choices. This appears to be the result of the different ages at which maximum prejudice scores occurred. At 7 years boys displayed a sudden rise in scores, whilst girls' scores remained relatively low until the age of 8 years.

After the age of 8 years there would appear to be a decrease in prejudice scores (chi-square values decreasing from 98·02 at 8 years to 26·82 at 10 years; see Table 1). This decrease in scores may be partly due to the concentration of coloured immigrant pupils in the schools from which these subjects were taken. This is explained below.

The schools were separated into those with immigrant pupil concentrations of 12 per cent and less, and those with immigrant pupil concentrations of 22 per cent or more. There were no schools with concentrations between 12 and 22 per cent. Of the two play groups one had no immigrant children and the other had 15 per cent.

Table 2. *Analysis of positive statements attributed to shaded and non-shaded figures by sex*

Age-group	Chi-square (with Yates' correction): positive statements *v.* sex	*P*
3–3·11	0·0	n.s.
4–4·11	0·62	n.s.
5–5·11	0·02	n.s.
6–6·11	0·55	n.s.
7–7·11	12·66	0·001
8–8·11	0·08	n.s.
9–9·11	0·0	n.s.
10–10·11	1·31	n.s.

Table 3. *Percentage of 'positive to non-shaded' responses in establishments with high and low immigrant concentrations*

Age-group	Immigrant concentration		Chi-square (with Yates' correction)	*P*
	Low	High		
3–3·11	50·0	51·9	0·002	n.s.
4–4·11	53·0	48·0	0·16	n.s.
6–6·11	66·4	46·9	4·76	0·05
9–9·11	79·0	70·8	0·56	n.s.
10–10·11	68·9	50·0	5·06	0·05

Table 3 shows the percentage of positive statements attributed to the non-shaded figure in these types of schools, and Table 4 the percentage of negative statements attributed to the shaded figure.

In the primary school age-range there were only three age-groups which included subjects from high immigrant concentration schools. Tables 3 and 4 indicate that in these age-groups there were substantial differences in the ways subjects responded. Subjects in schools with high concentrations of immigrants scored in less prejudiced ways. These differences were significant for the 6–6·11 years and 10–10·11 years groups ($P < 0{\cdot}05$).

The apparent decrease in prejudice in the 9–9·11 years and 10–10·11 years groups is partly attributable to the relatively large proportion of high immigrant schools in these samples. The differential effect of immigrant populations on the results has been discussed above. When the high immigrant schools sample is removed, the remaining sample displays no significant decrease in prejudice in the 9–9·11 years range and a much smaller decrease in the 10–10·11 years range.

Tables 3 and 4 also show the scores obtained by subjects in the play groups. As can be seen, there is little difference between these. In the 4–4·11 years group the subjects from the low immigrant concentration play group scored in a slightly more prejudiced way, but this difference was not significant. The similar increase in attribution of negative statements to the non-shaded figure by 3-year-olds from the high immigrant play group was not significant.

Table 4. *Percentage of 'negative to shaded' responses in establishments with high and low immigrant concentrations*

Age-group	Immigrant concentration		Chi-square (with Yates' correction)	*P*
	Low	High		
3–3·11	64·3	40·4	3·21	n.s.
4–4·11	52·0	48·0	0·07	n.s.
6–6·11	63·3	45·8	3·96	0·05
9–9·11	75·0	65·3	1·28	n.s.
10–10·11	70·8	53·6	4·28	0·05

DISCUSSION

This study gives results of children's reactions to figures with 'white' (non-shaded) and 'grey' (shaded) skins. Therefore any conclusions which are drawn cannot claim more than an awareness of these contingencies, but on the basis of the Williams & Edwards study (1969) one might tentatively infer a link between the findings presented here and children's racial attitudes.

At the ages of 3 and 4 years children tended to select in a random fashion, but from the age of 5 onwards the attribution of statements to figures was increasingly predictable. Radke *et al.* (1950) also found very high levels of prejudice among Caucasian children with a very similar test.

After correction for the presence of children from high immigrant concentration schools there was still a small but significant decrease in prejudice in the 10-year-olds. No explanation is offered for this, but it should be noted that Horowitz & Horowitz (1938) and Radke *et al.* both commented upon similar decreases in prejudice scores around 10 years of age. Whilst the slight decline noted in this study may be indicative of a growing awareness of less prejudiced attitudes amongst the older children, this would need substantiating.

From this study there is evidence to suggest that contact with coloured immigrant children leads to a decrease in the level of prejudice displayed by Caucasian children. It might be inferred that the reduction of prejudice choice to shaded and non-shaded figures is indicative of a modification of racial concept attitudes. This would add to the body of evidence recommending the cessation of racial segregation in the schools or the community at large. 'There is evidence for stating that if the social realities present within the situational context permit the experiencing of equal status contacts, the frequency of attitudes expressing inter-racial antagonisms will be reduced' (Winder, 1955, p. 101).

Children with no close contacts with immigrants tend to rely on stereotypes derived from hearsay, or from atypical situations. Such stereotyping, and its associated prejudices, is diminished by close contact with immigrants, since there is then an opportunity to form evaluative judgements which are truly representative of the racial group.

No significant differences were found between the levels of prejudice exhibited by males and females. This supports the findings of Horowitz & Horowitz (1938). Sex differences were reported by Landreth & Johnson (1953), though these were only significant in relation to socio-economic status.

REFERENCES

Clark, K. B. & Clark, M. P. (1947). Racial identification and preference in Negro children. In T. M. Newcomb & E. Hartley (eds.), *Readings in Social Psychology*. New York: Holt.

Harbin, S. P. & Williams, J. E. (1966). Conditioning of colour connotations. *Percept. mot. Skills* **22**, 217–218.

Horowitz, E. L. & Horowitz, R. E. (1938). Development of social attitudes in children. *Sociometry* **1**, 301–338.

Isaacs, H. R. (1963). *The New World of the Negro American*. New York: Day.

Landreth, C. & Johnson, B. C. (1953). Young children's responses to a picture inset test designed to reveal reactions to persons of different skin colours. *Child Dev.* **24**, 63–80.

Radke, M. J., Sutherland, J. & Rosenberg, P. (1950). Racial attitudes of children. *Sociometry* **13**, 154–171.

Renninger, C. A. & Williams, J. E. (1966). Black–white colour connotations and racial awareness in pre-school children. *Percept. mot. Skills* **22**, 771–785.

STEVENSON, H. W. & STEWART, E. C. (1958). A developmental study of racial awareness in young children. *Child Dev.* **29**, 399–409.

WILLIAMS, J. E. (1964). Connotations of colour names among Negroes and Caucasians. *Percept. mot. Skills* **18**, 721–731.

WILLIAMS, J. E. (1966). Connotations of racial concepts and colour names. *J. Person. soc. Psychol.* **3**, 531–540.

WILLIAMS, J. E. (1969). Individual differences in colour-name and connotations as related to measures of racial attitude. *Percept. mot. Skills* **29**, 383–386.

WILLIAMS, J. E. & CARTER, D. J. (1967). Connotations of racial concepts and colour names in Germany. *J. soc. Psychol.* **72**, 19–26.

WILLIAMS, J. E. & EDWARDS, C. (1969). An exploratory study of the modification of colour and racial concept attitude in pre-school children. *Child Dev.* **40**, 737–750.

WILLIAMS, J. E. & ROBERSON, J. K. (1967). A method of assessing racial attitudes in pre-school children. *Educ. psychol. Measur.* **27**, 671–689.

WINDER, A. E. (1955). White attitudes towards Negro–White interaction in an area of changing social composition. *J. soc. Psychol.* **41**, 85–102.

Manuscript received 8 *July* 1970

Revised manuscript received 10 *November* 1970

Br. J. soc. clin. Psychol. (1971), 10, pp. 313–319
Printed in Great Britain

Introversion, Extraversion and Social Perception

BY STEVEN R. BROWN

Department of Political Science, Kent State University

AND CLYDE HENDRICK

Department of Psychology, Kent State University

Three studies are conducted on the subjective perceptions of extraverts and introverts, using *Q* technique and factor analysis. In the first two studies, intensive analyses indicate that both introverts and extraverts have extraverted aspirations. These results are confirmed and extended in the third study with a larger sample of subjects. Both introverts and extraverts have veridical perceptions of actual private self and actual social self; however, the ideal conceptions of both personality types tend to be extraverted in nature. Introverts express their potential private self, potential social self and ideal leader perceptions as those of an extravert; similarly, extraverted subjects hold extraverted ideals. Extraverts are highly consistent in their perceptions of actual and ideal selves, but introverts have ideals incongruent with perceived actual selves. Implications for political theory and behaviour, interpersonal perception and attraction are discussed.

During the past few years, Jung's (1923) classical theory of introversion–extraversion has received increasing attention, primarily by Eysenck and his students. Introversion–extraversion (E–I) has been conceptualized as a personality trait dimension, and a successful measure of this trait has been devised (Eysenck, 1962). A great deal of research has been carried out centring on the relationships between this trait and other traits, including the relationship of E–I to anxiety, conditionability, toughmindedness, aggression, mental deficiency, paranoia, persuasibility and auditory threshold, to name just a few. A recent summary of work along these and other lines is to be found in Eysenck & Eysenck (1969, especially Table 11.8, pp. 131–2).

Although modern research has tended to focus on objective trait relationships, Jung was primarily interested in underlying subjective dynamics. One of his propositions, for example, was that extraverts and introverts have very different perceptual perspectives of their social worlds. In particular, introverts and extraverts may misperceive one another since 'the introvert makes the mistake of always wanting to relate action to the subjective psychology of the extravert, while the extravert can only conceive the inner mental life as a product of external circumstances' (Jung, 1923, p. 205).

The present research was concerned with exploring similarities and differences in the subjective experiences of introverts and extraverts. *Q* technique and factor analysis, previously employed in research on the Jungian typology (e.g. Stephenson, 1939; Gorlow *et al.*, 1966), were the tools used in this exploration. Several aspects of interpersonal perception as a function of E–I were investigated, including actual and desired perceptions of self, perceptions of introverts by extraverts, and perceptions of extraverts by introverts. In addition, special emphasis was placed on the differential perception of leadership qualities.

Three studies relevant to these questions are reported. The first two studies were preliminary in nature, depending upon intensive analysis—as opposed to extensive analysis (Shapiro, 1961)—of a few cases. The results of the first two studies were then given an extended test in the third study.

I

Students in a graduate seminar in political personality were given Eysenck's (1962) Maudsley Personality Inventory (MPI) and the three subjects with the highest extraversion scores (extraverts) and the three subjects with the lowest extraversion scores (introverts) were chosen for intensive analysis. The E scores of the extraverted and introverted subjects chosen were in the upper and lower 30 per cent of the MPI range, respectively. The subject with the highest E score (the model extravert) and the subject with the lowest score (the model introvert) were asked to provide their self-descriptions with a Maudsley *Q* sort, which contained the MPI items. The *Q* sample contained 24 items from the E–I component of the MPI (12 introvert items, 12 extravert items) and 24 items from the neuroticism component (12 anxiety items, 12 stability items). Some statements were reworded to provide equal numbers of the four types of items. The *Q* sort continuum ranged from +4 to −4 in a forced, quasinormal distribution.

The other two introverts and other two extraverts were instructed to describe the two model subjects using the same *Q* sample, i.e. the two other extraverted subjects were instructed to use the MPI *Q* sort to describe what the model extravert was like, and then to describe what the model introvert was like; the two other introverted subjects likewise described the models. The resulting 10 *Q* sorts were correlated and the correlation matrix produced two factors. Factor 1 contained only one defining variable, the model introvert's self-description. The *Q*-sort attempts to describe the model introvert by the other introverts and the extraverts, all of whom were the subject's associates, failed to load significantly on his factor. Factor 2 was defined by the model extravert's self-description and all of his introverted and extraverted friends' successful attempts to describe him. In other words, the extravert was highly visible and was correctly perceived both by fellow extraverts and by introverts, the latter result being contrary to Jung's postulation. The introvert, on the other hand, was misperceived not only by extraverts, as suggested by Jung, but by other introverts as well.

II

In an effort to discover if there was something the introverted subject might be doing which was preventing others from perceiving him correctly, and/or if there was something the extraverted subject was doing which was making him so visible, both subjects used in the first study were asked to provide four *Q* sorts (again using the Maudsley *Q* sample) according to four conditions of instruction: (1) *actual private self*, 'the kind of person I am'; (2) *actual social self*, 'the kind of person other people think I am'; (3) *potential private self*, 'the kind of person I would like to be'; and (4) *potential social self*, 'the kind of person I would like other people to think I am'. These representative probes of the subjects' subjective behavioural capabilities were not made with any *a priori* expectation as to what would result, but for purposes of broadening the inductive base from which to draw inferences which hopefully would bear on introversion and extraversion more generally.

The eight *Q* sorts—four each from the introvert and extravert—were factored, again resulting in two factors. The extravert's four *Q* sorts were significant and defining on factor 1, which indicates that his extravertish self-description is also how he thinks others regard him, is also what he regards as personally ideal, and is also descriptive of what he would like others to think he is like. The introvert's *actual private self* and *actual social self* were introverted in content and were the sole defining variables on factor 2. However, the introvert's *potential private self* and his *potential social self* were purely loaded on the extraverted factor 1. From a factor-analytic point of view, while the extravert's self system is homogeneous, the introvert's is heterogeneous, i.e. the introvert's ideals (private and social) are extraverted in quality.

The results obtained provide a possible explanation as to why the introverted subject was incorrectly perceived in the first study. As an introverted type (as defined by the MPI), he described himself as such in his *actual private self Q* sort; however, his aspirations (*potential private self* and *potential social self*) appear extraverted, i.e. he appears to want others to see him as other than what he is. Although he fails to transmit an extraverted image, his 'extraverted act' apparently succeeds in providing sufficient interpersonal noise as to make accurate perception of him difficult. The relative actual–potential congruence by the extraverted subject tends to minimize behavioural inconsistencies and renders accurate perception more probable.

III

Subjects and procedure. The subjects were university seniors in political science. Each was administered the MPI, and 10 subjects with high E scores (extraverts) and 10 subjects with low E scores (introverts) were retained for further testing. All extravert subjects were in the upper 30 per cent of the MPI scale; all introvert subjects were in the lower 30 per cent. The same Maudsley *Q* sample used in the first two studies was employed.

Subjects were instructed to provide four *Q* sorts under the same conditions of instruction as in the second study: (1) *actual private self*, (2) *actual social self*, (3) *potential private self*, and (4) *potential social self*. On the basis of the first two studies, consideration was given to the possibility that the introvert might have a tendency to view as extraverted many future and/or socially remote and desirable objects (e.g. potential selves) while continuing to view as introverted the more intimate and immediate objects (e.g. actual selves). As an extension to the previous results, therefore, each subject was asked to provide his conception of a remote, and at the same time socially relevant, object, (5) the *ideal leader*.

With 10 introverts and 10 extraverts and five *Q* sorts from each, 100 *Q* sorts resulted, which were correlated and factored using the principal axis method. Four factors were extracted and rotated by varimax criteria to a position in simple structure. The fourth factor contained significant loadings for only two of the 100 variables and was dropped from consideration.

RESULTS

Factor scores. Factor scores associated with the MPI scale items in each factor indicate factor 1 is the prototype of the typical extravert while factor 2 is exemplary of the non-neurotic introvert. What has been referred to as the 'dual nature of extraversion' (Eysenck & Eysenck, 1969, pp. 141–9) is supported to some degree by factor 3 which is a more stable (non-neurotic) variation of factor 1. The factor-score mean for the extraversion items is higher than the mean for the introversion items in factors 1 and 3 ($P<0·01$), and introversion is higher than extraversion in factor 2 ($P<0·01$). As far as the theoretical content of the factors is concerned, therefore, factor 1 may be regarded as typically extraverted, factor 2 as typically introverted, and factor 3 as extraverted and non-neurotic.

Factor loadings. The subjects were divided by personality type (E–I) and each provided Q sorts for the five psychosocial conceptions (*potential social self*, *ideal leader*, etc.). After transforming the factor loadings to Fisher's Z, each factor was analysed as a repeated measures design with one between effect (personality type) and one within effect (psychosocial conceptions). The results are shown in Table 1.

Table 1. *Analysis of variance of factor loadings*

		Factor F ratios		
Source	D.F.	1	2	3
Between subjects	19			
Personality (A)	1	13·309**	2·880****	0·199
Error 1	18			
Within subjects	80			
Conceptions (B)	4	6·194***	2·111*	37·885***
A × B	4	14·907***	1·494	2·060*
Error 2	72			
Total	99			

* $P<0·10$. ** $P<0·01$. *** $P<0·001$. **** $P<0·1036$.

The highly significant interaction on factor 1 (which is extraverted in terms of factor scores) is due primarily to the behaviour of the introvert subjects, whose factor loadings for *actual private self* and *actual social self* conceptions average $-0·16$ and $-0·04$, respectively, but whose factor loadings for *potential private self*, *potential social self*, and *ideal leader* conceptions average 0·27, 0·38, and 0·30, respectively. All five conceptions for the extraverted subjects, however, averaged positively, as indicated in Table 2. The small but significant interaction for factor 3 (which is extraverted and stable in content) is due to divergences in conceptions for introverts and extraverts. Table 2 indicates that for both introverts and extraverts, the idealized conceptions (*potential private self*, *potential social self*, and *ideal leader*) have high-positive average factor loadings, while the *actual private self* and *actual social self* conceptions are lower for introverts than for extraverts.

Factor 2, which is introverted in content, is problematical by virtue of the low average factor loadings, even for the introverted subjects' *actual private self* conception. This is due in part to the fact that some of the introverts (as defined by the

Table 2. *Factor loading means*
(Based on Fisher's Z transformation)

	Typical extravert 1	Typical introvert 2	Stable extravert 3
Extraverts			
Actual private self	0·55	−0·15	0·28
Actual social self	0·47	−0·06	0·37
Potential private self	0·41	0·02	0·61
Potential social self	0·52	−0·10	0·52
Ideal leader	0·24	0·10	0·72
Introverts			
Actual private self	−0·16	0·09	0·12
Actual social self	−0·04	0·20	0·24
Potential private self	0·27	0·10	0·64
Potential social self	0·38	0·08	0·62
Ideal leader	0·30	0·13	0·69

MPI) described themselves (using the Maudsley Q sample) in such a way as to be bipolar to the extraverts on factor 1, whereas other introverts described themselves in such a way as to be orthogonal to the extraverts, i.e. as positive and purely saturated on factor 2, suggesting a dual nature to introversion as well as extraversion. Therefore, whereas factor 2 contains *only* introverts, it does not contain *all* introverts and the introvert mean factor loading is reduced as a consequence. Nevertheless, the average factor loading for the introvert *actual private self* conception exceeds that for the extravert ($P < 0{\cdot}01$), and the average introvert factor loading is higher than the average extravert factor loading for each of the other four conceptions, although this difference is not always significant.

DISCUSSION

The third study confirms the results of the first two. Introverts' ideals and aspirations tend to be extraverted in nature:* whereas introverted subjects' *actual selves* (private and social) were either on the introverted factor 2 or negative on the extraverted factor 1, their *potential selves* (private and social) were positive on both of the two extraverted factors 1 and 3. Both introverts and extraverts choose factor 3 as the expression of their conception of the *ideal leader*, which implies that there is an expectancy for the *ideal leader* to be a stable extravert. The results have a number of implications. The results for leadership have implications for empirical political theory. The results for leadership as well as the other results are also relevant to the general area of interpersonal perception.

---

* Research of this kind must always carry the proviso that the results may not generalize across national boundaries. Just as E–I norms differ somewhat from country to country, so may social perceptions. But by the same token, just as introversion and extraversion appear to emerge in some systematic form in many countries (Eysenck & Eysenck, 1969, pp. 90–6), so, too, would we expect the perceptions and subjective dynamics supported by these personality types to make systematic appearances.

Political theory. Prior to Eysenck's (1954) pioneering study, McConaughy (1950) found a group of political leaders to be more extraverted than normals (i.e. the groups on which the tests had been standardized). Hennessy (1959), on the other hand, suggests that while the image of the politician is one of outgoing extraversion, the image is probably not grounded in fact. That the leader in general is at least expected to be extraverted is apparent in the above results, and there is evidence that practical politicians and political image-makers, at least in the United States, are aware of this expectation. McCaffrey (1962), for example, points to a narcissistic *prima donna* characteristic qualifying men for political candidacy, and in constructing his composite 'winning candidate' includes such extraverted qualities as stamina, oratory, aggressiveness, friendliness, cordiality, 'the sincere enthusiasm of a little boy,' winning humour and self-effacing wit.

Whether or not political leaders are more extraverted than the average person remains open to further investigation. However, the above results suggest that the extravert is best equipped to gain mass support since his personality style is one with which both extraverts and introverts can identify: he is the ideal for both. Although the extravert tends in general to be more socially adjusted, he nevertheless has certain possibly negative attributes which under specified conditions might have political consequences. As Eysenck (1954), for example, has noted, 'the extravert is prone to disorders like hysteria and psychopathy, i.e. asocial illnesses in which moral rules tend to be disregarded' (p. 175).

The functional relevance to a political system of one political personality type or another is dependent on situational requirements which determine not only policy direction (e.g. radical or conservative) but also the style in which such policies will be executed (e.g. by evolution or revolution). Periods of social uncertainty place a premium on the manifestly self-assured, narcissistic individual with whom the mass can identify and whose behavioural reinforcement is contingent upon 'the emotional response of the public' (Lasswell, 1960, p. 78). Such self-accepting behaviour is more typical of the extravert, as is the requisite other-oriented behaviour. Periods of relative calm and consensus, on the other hand, place a premium on the controlled, introspective, more introverted and administrative type, whose reinforcements are concrete and intimate rather than abstract and remote. Changing social circumstances influence a society's elite recruitment patterns and bring into prominence certain personality types with clusters of skills which have implications for the maintenance of the political system.

Interpersonal perception. The above results suggest the introvert is capable of a wider range of subjective behaviour by virtue of his extraverted aspirations. The preliminary study indicated a lower 'perceptual visibility' for introverts, both because of the possible relative insensitivity of the extravert to interpersonal cues and because of the discrepancy between what the introvert is actually like and what he may be trying to act like.

In terms of interpersonal attraction, the extravert is said to be scornful of his opposite, for he tends to regard the introvert as 'either a conceited egoist or a fantastic doctrinaire' (Jung, p. 476). The extravert therefore will have a tendency to be interpersonally attracted to others on the basis of their perceived similarity to himself. The introvert, however, does not react negatively to the extravert and so his friendship circle is potentially more heterogeneous, including both introverts and extraverts. Recent research (Hendrick & Brown, 1971) indicates that the introvert under certain conditions is highly attracted to extraverts, even more so than to other introverts.

REFERENCES

EYSENCK, H. J. (1954). *The Psychology of Politics*. London: Routledge & Kegan Paul.

EYSENCK, H. J. (1962). *The Maudsley Personality Inventory*. San Diego, Calif.: Educational & Industrial Testing Service.

EYSENCK, H. J. & EYSENCK, S. B. G. (1969). *Personality Structure and Measurement*. London: Routledge & Kegan Paul.

GORLOW, L., SIMONSON, N. R. & KRAUSS, H. (1966). An empirical investigation of the Jungian typology. *Br. J. soc. clin. Psychol.* **5**, 108–117.

HENDRICK, C. & BROWN, S. R. (1971). Introversion, extraversion, and interpersonal attraction. *J. Person. soc. Psychol.* (in press).

HENNESSY, B. (1959). Politicals and apoliticals: some measurements of personality traits. *Midw. J. pol. Sci.* **3**, 336–355.

JUNG, C. G. (1923). *Psychological Types*. London: Routledge & Kegan Paul.

LASSWELL, H. D. (1960). *Psychopathology and Politics*. New York: Viking Press.

MCCAFFREY, M. (1962). *Advertising Wins Elections*. Minneapolis, Minn.: Gilbert.

MCCONAUGHY, J. B. (1950). Certain personality factors of state legislators in South Carolina. *Am. pol. sci. Rev.* **44**, 897–903.

SHAPIRO, M. B. (1961). The single case in fundamental clinical psychological research. *Br. J. med. Psychol.* **34**, 255–262.

STEPHENSON, W. (1939). Methodological consideration of Jung's typology. *J. ment. Sci.* **85**, 185–205.

Manuscript received 10 *July* 1970

Revised manuscript received 9 *November* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 320–331
Printed in Great Britain

Judgement of Direction of Gaze: An Interpretation of Discrepant Results

By IAN VINE

Department of Psychology, University of Bristol

Recent research on non-verbal behaviour in face-to-face social interaction has suggested that direction of gaze may be an important variable. However, investigators have paid little direct attention to the reliability and validity problems of this measure, relying largely on results by Gibson & Pick (1963) and Cline (1967) which showed good discrimination in an artificial experimental setting. Some recent studies have thrown doubt on the assumption of high validity of observer judgements of direction of gaze. This in turn throws doubt on results of studies using direction of gaze as a variable. Analysis of reliability and validity studies suggests that the discrepant results may be reconciled. It is argued that in natural interactions it may be sufficient if observers can accurately discriminate between gazes at the face of another and gazes directed away from the face, an easier task than in several laboratory studies. Nevertheless, in view of the possible bias effects involved, caution is needed in interpreting results involving assessments of gaze directions.

During social encounters we attend to our fellow-interactors both aurally and visually, and the importance of the continuous monitoring of both types of input may be considerable. Recent reviews of studies where the amounts of other-directed gazing during interactions are monitored indicate that this gazing can have important functions (Argyle & Kendon, 1967; von Cranach, 1971*b*; Duncan, 1969, Vine, 1970). Looking at another may indicate attention, liking or threat (depending on the duration of the look, and accompanying signals) and so on, while patterns of looking and not looking during conversations are intimately linked to the regulation of information exchange. Individual differences in typical amounts of looking can also be related to various personality characteristics or psychiatric states. A variety of aspects of our social behaviours and interpersonal evaluations may thus depend on how and where we direct our gaze during interactions, and how well we detect this in others.

This latter question raises methodological problems for the quantitative study of gazing phenomena. If judgement of gaze direction (GD) is typically poor, then the results of a rapidly growing body of research are in doubt, for as yet this has relied on observer judgements in monitoring the gazing variables. Some recent evidence on the validity of judging direction of gaze suggests that substantial systematic biases may affect judgements and that inter-observer reliability can be low (von Cranach, 1971*a*, *b*; Ellgring, 1971; Stephenson & Rutter, 1970). The status of findings from GD studies thus depends on how well the discrepant results on validity and reliability of judgements can be resolved.

Most GD studies have not reported on formal validity or reliability assessments of observer judgements, often relying on the results of Gibson & Pick (1963) and of Cline (1967) in assuming valid discrimination between a gaze directed at the eyes of the other and gazes directed elsewhere. At best they have reported high inter-observer agreement in such judgements. But agreement alone does not validate gaze discrimination by the observers, and there is now some doubt as to just what are the implications of the above two studies in social contexts. If the studies of GD during social interaction are to be convincing, we must establish that it can be judged accurately. It will be argued here that existing data can be reconciled with this conclusion if we consider the factors which are really appropriate to their evaluation, in particular, whether or not gazing behaviour is determined by the experimenter rather than the free choice of the interactors.

PSYCHOPHYSICAL STUDIES OF JUDGEMENTS OF GAZE DIRECTION

The experiments by Gibson & Pick (1963) and Cline (1967) measured visual acuity for discrimination between directions of different fixations by a sender, corresponding to targets on or near the face of the receiver making the judgements. In the basic experimental paradigm the sender fixates various targets in random order, and on each occasion the receiver makes a judgement, either saying whether a particular fixation is directed at his own eyes (in an eye-gaze, or EG) or elsewhere, or specifying the exact judged location of the target fixated. Effects of the orientation of the sender's head have also been investigated for head directions (HDs) independent of or coincident with GDs. Recently Anstis *et al.* (1969) went even further in presenting the sender via a television screen, the orientation of which could itself be varied. The principal features and results of these studies are listed in Table 1.

The findings of the above investigations can be summarized as follows. (1) Gibson & Pick concluded that GD judgements are likely to approach the human acuity limits in terms of discrimination, but where the head is turned away from the axis of GD, deviations of the GD itself from the sender–receiver axis will be underestimated by an amount dependent on the HD; (2) Cline found accurate judgement of actual EGs, but poorer discrimination of other GDs, and confirmed that there is underestimation of deviations when the head is turned independently; (3) Anstis *et al.* found that the amount by which a GD is misjudged is proportional to its deviation from the sender–receiver axis, and the deviation is overestimated. Turning the head increased this effect, but yielded constant errors of underestimation as in the earlier studies.

In view of the methodological differences between these studies the differences between their results are not easy to interpret, a point to which we shall return shortly. However, they do suggest that discrimination of GD is rather good, at least inasmuch as actual EGs *are* likely to be perceived as such. They also agree in indicating that if the HD and GD are not coincident this yields a systematic bias which in itself tends to lead to underestimates of deviation of the GD from the sender–receiver axis. In the study by Anstis *et al.* (1969) further experiments not mentioned in Table 1 investigated the possible reasons for this underestimation, and for their own systematic overestimation effect. Instead of using a human subject they used an artificial eyeball set in a variable diaphragm, and were able to show that both effects could be explained geometrically. It appears that GD is judged on the

Table 1. *Psychophysical studies of judgements of gaze direction*

Author	Sender-receiver distance	Targets (locations relative to bridge of receiver's nose	Head orientations (relative to 'nose-to-nose' sender-receiver axis)	Measures used and main results
Gibson & Pick (1963)	200 cm	Central, also 10 cm, 20 cm, 30 cm to either side drawn on scale behind receiver's head) (angles are 2·9°, 5·7°, 8·5° approx.)	Central, also 30° to either side, for each target	Binary EG *v.* non-EG judgements, yielding an EG frequency plot across fixation targets for any given GD plus HD combination. Mean value indicates 'constant error' bias (CE), standard deviation used as discrimination threshold.
				Threshold independent of GD/HD and corresponding to typical Snellen acuity, being 1·10 min of arc. Head-turn leads to constant underestimate of GD deviation from sender–receiver axis (more false positive EGs)
Cline (1967)	122 cm (Receiver sees mirror-image of sender with target board superimposed)	Central, also 2°, 8°, 12° to either side and up and down*	Central	Judging which of 65 points on target board (which included the 13 actual targets) was fixated. Threshold and CE measures as above separating both vertical and horizontal error components.
				Central fixation threshold 1·55 cm horizontal, 2·70 cm vertical (former corresponds to a GD at 0·75° to sender–receiver axis, and represents 0·51 min of arc acuity of receiver), negligible CE. Larger thresholds and CEs for other GDs (up to 8·59 cm and 8·36 cm respectively).
		Central, also 4°, 10° to either side	Central, also 30° to right, or head coincident with GD, for each target	Threshold for central target now 2·85 cm horizontally, but 5·29 cm (with CE of 3·00 cm) with independent head turn. All conditions gave larger thresholds for other GDs (up to 6·75 cm), but while CEs were up to 5·26 with the central or 30° HD, they were lower with coincident turns (up to 1·40 cm). Non-coincident turns thus gave underestimation biases in judging deviation of GDs from sender–receiver axis.

---

[*] Cline's Figs. 2*A*, 2B appear to show 4° as his minimum deviation. This is one of several ambiguities in his report.

Anstis et al. (*1969*)	*84 cm (chosen to give life-size image with TV presentation)*	*Central, also 5°, 10°, 15°, 20° to either side and* up or down (fixations were actually on a scale midway between sender and receiver, also used by receiver to indicate judgements)	*Central, also 30° to either side, for horizontal* targets; central for vertical targets Central, also 30° to either side, but image presented on TV screen itself central or 30° to either side	*Moving a pointer to the judged fixation position on the scale*; measures used to obtain regression equations of the form: judged deviation = *m* (actual deviation) + constant error. With central target judgement almost veridical, also for vertical targets. With horizontal targets judged GD deviations from sender–receiver axis were from 50 per cent (head central) to 69 per cent overestimated. Head-turn also increased CE as high as 3·85°, this being an underestimation bias reducing the systematic effect. Overestimation of deviations higher, up to 87° with screen and head turned. Turning the screen itself introduced an overestimation CE of about half the extent of the opposite head-turn CE.
Ellgring (1971)	80 cm and 200 cm	Central, also 3 cm, 6 cm to either side and up and down (angles at 80 cm are 2·1° and 4·3° approx., at 200 cm 0·9° and 1·7° approx.)	Central	Judging which target was fixated, yielding frequency and 'information' indices. Percentage correct judgements decreased with distance from 50·1 to 29·9 per cent. For all targets judgements were greater than chance. While frequency of correct judgement was slightly greater for points around the eyes, relative to number of times a target was named by a judge accuracy was greatest for peripheral targets, and lowest for the central point.
Krüger & Hückstedt (see von Cranach, 1971*a*)	150 cm and 300 cm (also observers judged from various distances and orientations to sender–receiver axis)	Central, also 6 cm, 16 cm, 26 cm to either side, or 3 cm, 13 cm, 23 cm up, 3 cm, 13 cm, 23 cm, 33 cm down (6 cm at 150 cm gives 2·3°, at 300 cm 1·1°)	Central, also 20° to either side	Binary FG *v.* non-FG judgements, yielding frequency indices. Fixations of the three inner targets actually on the face were usually recognized as such by receivers, but non-FGs were often judged as FGs, and FG response was higher at the greater distance. Differentiation of targets by observers was poor, especially in oblique observation orientations. Observers' judgements depended mainly on HD.

Table 1—*continued*

Author	Sender-receiver distance	Targets (locations relative to bridge of receiver's nose	Head orientations (relative to 'nose-to-nose' sender-receiver axis	Measures used and main results
von Cranach & Hückstedt (unpublished)	As above	As above, plus gaze movement preceding fixation	Central	Measures as above. Gaze judgement was not improved by movement, but movement towards the face yielded more false positives where the subsequent fixation was actually a non-FG.
von Cranach, Hückstedt, Schmid & Vogel (unpublished)	As above	As above, with and without gaze movement	Central, also 15°, 30° to either side, and head movement preceding fixation	Measures as above. Again at the greater distance, discrimination of FG *v.* non-FG was worse, so that HD now mainly determined judgements. Head movement tended to reduce the number of FG responses; its extent affected only observers and at 300 cm. Observers again discriminated worse than did receivers. In general, discrimination is worse with both movements.
Ellgring (see von Cranach & Ellgring, 1971)	300 cm	Central, also 20 cm, 40 cm to either side (angles 3·8° and 7·6°)	Central	Judging which target was fixated, yielding frequency indices. Total of 100 judgements elicited for each target, over about 1 hour, with feedback as to correct targets. Reinforced practice led to no improvement for central target, but improvement for more peripheral ones, eliminating the bias towards citing more central targets in judgements.

basis of the perceived position of the iris in the visible part of the eye, ignoring the fact that the latter is only part of the surface of a larger sphere. This leads to overestimation of movements. Turning the head affects the relative amounts of white seen on either side of the iris in such a way as to yield underestimation, by similar principles. Thus both eye position and head orientation yield cues for judgement of GD. (Judgement of *vertical* deviations of gazes fron the sender–receiver axis would be similarly cued by eye and head, but the relationships differ. Discrimination is in fact good according to Cline (1967) and to Anstis *et al.* (1969), and the latter study found it not susceptible to overestimation.)

If systematic overestimation is, as appears, an overriding effect, one would expect participants in social interactions to be adept at discerning whether or not they were receiving EGs, as any other gazes would be perceived as more deviant from EGs than they in fact were, and confusion in the simple EG *v.* non-EG judgements would be low. However, recent data from von Cranach and his associates in Munich suggest that there is *no* special sensitivity to EGs, and in typical situations the systematic overestimation found by Anstis *et al.* may be absent. They have also gone further in considering judgements of GD not only by the receivers, but also by observers viewing the sender's fixations of the receiver. A number of studies by the group are discussed by von Cranach (1971*a*, *b*), and these also are summarized in Table 1. Their main conclusions are as follows: (1) discrimination between target points in and around the face is moderate at close sender–receiver distances for the receiver, but is fairly poor even at 300 cm (in fact, Krüger & Hückstedt first performed the study which Ellgring, 1971, replicated, and they found less than chance accuracy at this distance); (2) errors show a trend towards underestimation of the GD discrepancy from the sender–receiver axis, which increases with distance; (3) when judgement is simplified into discrimination of gazes within the face or outside the face (*face-gazes*, or FGs *v.* non-FGs), most actual FGs are detected, but again there are false positives, and the underestimation of GD deviation from the sender–receiver axis increases with distance; (4) indirect HDs reduce accurate discrimination and have more influence on judgement at greater distances, especially for observers; (5) additional cues of gaze or head movement prior to the fixation to be judged reduce accuracy; (6) the observer's judgement depends more on HD than does the receiver's, and accuracy is lower, even when the observer stands almost directly behind the receiver.

While the Munich findings cannot be translated directly into statements on the accuracy of estimating gaze deviations from the sender–receiver axis, they clearly indicate that discrimination is far from perfect, even in the best conditions, and above the threshold GD deviation of about 1° that Cline (1967) found. Another finding of the studies was that between-judge variance was appreciably (50 per cent) due to variations in visual acuity. If acuity is a limiting factor one would expect discrimination to deteriorate with distance, but the poor discrimination even between FGs and non-FGs, which is barely influenced by actual GD in the worst conditions, is surprising.

The conflict between the Munich results and the other studies is not easily resolved, especially as the methodologies of experiments have differed considerably.

If we compare the study by Ellgring (1971) with that of Cline (1967), the gaps between targets were of a smaller angle, especially at the larger distance, so finer discrimination was actually demanded, a point which is masked by Cline's use of standard deviations in deriving a threshold value. (Serious doubts can in fact be raised about the use of such measures; see von Cranach & Ellgring, 1971.) Given the more difficult task, we might expect poorer accuracy, but also more uncertainty of the receiver's judgements. Ellgring's data showed an underestimation bias which may be partly explained by this uncertainty. Such a bias would give a smaller average error magnitude than would random choice where judgements were very difficult. This does not, however, explain why Cline found his threshold lowest for actual EGs.

Ellgring's (1971) study also conflicts with that of Anstis *et al.* (1969), who found a systematic overestimation effect. Ellgring has recently shown that practice can be important in judging GDs (von Cranach & Ellgring, 1971). When receivers were given feedback as to correct fixation targets, they improved over a series of 500 judgements, *except* in judging actual EGs. Improvement increased with increasing GD deviations from the sender–receiver axis. Variance of judgements also decreased, and the underestimation bias again found at the start of the trials was eliminated by the end of the period. Ellgring considers that the data suggest ability to detect actual EGs is already overlearned, while training is required for accurate detection of other GDs (see also Table 1).

The targets in the above study were separated by angles of about 3·8°, much more than in his previous study, and so the explanation of the underestimation bias in terms of a strategy appropriate to high uncertainty of judgements is less convincing here. However, comparing this study with that of Anstis *et al.* (1969), there are reasons why uncertainty should be greater in the early trials than in those of Anstis *et al.* These authors used a matching technique involving interval scale judgements, and their range of GDs was ± 20°. Ellgring's receivers had to make multiple choices between five target possibilities, and the range of GDs was only ± 7·6°. With this method subjective confidence could well be lower, and underestimation more likely. It is also of note that Anstis *et al.*'s sender was fixating points not actually on the face of the receiver, but well in front of him. The effect of this odd situation is somewhat imponderable, but it may make their results incomparable with those of other authors. Most noteworthy is the large angular range used by Anstis *et al.*, which leads to the speculation that perhaps their systematic overestimation effect is only found at greater angles. Inspection of the graphs from which their regression equations were derived indicates some appreciable deviations from linearity. It is particularly noticeable that the first signs of overestimation of gaze deviations occur with the 10° target; the 5° target shows no such bias of judgements obtained. We may thus conclude, tentatively, that Anstis *et al.* may have been unjustified in assuming that the relation between actual and judged GD orientations is linear.

On the basis of the studies so far discussed, we may formulate two hypotheses regarding bias in GD judgements: (1) that conditions of high judgement uncertainty may, where multiple-choice decisions are demanded, lead to underestimation biases yielding too many EG responses; (2) where precise measures are used, and uncertainty is lower, overestimation biases may apply when GDs are angled substantially to the sender–receiver axis. Further investigations, which include the elicitation of measures of subjective uncertainty of judgements and sample more suitable ranges of angles, should test these hypotheses. In these respects the results of the psychophysical studies must leave us in some doubt as to whether GD judgements can, in appropriate conditions, be as accurate as human visual acuity would theoretically permit. Also, it is evident that accurate judgement of whether a gaze is actually an EG, or even an FG, must depend on the distance involved between sender and receiver or sender and observer. It is also evident that discrepancies between gaze and head orientations lead to errors, so that judging GD in natural social interactions, especially where there are also movements of eyes and head of the senders, will be much less accurate than the earliest experiments suggested.

To conclude this section, we may mention some further variables which may potentially affect accuracy of GD judgements or the validity of conclusions from the studies discussed here. In studies to date there has not been independent monitoring of actual target fixations,[*] so it has been necessary to assume that senders do fixate the points as instructed. There could be systematic bias on the part of a sender, or there could be random error. In fact, nystagmus movements would be expected to yield some fluctuations of fixations, and duration of fixations may thus be important for this and other reasons. Too short fixations will mean that the average location of the fixation corresponds less well with the intended target, and if target separations are small this might increase judgement difficulty and the receiver's uncertainty. Too long fixations might be subject to fixation drift due to fatigue. Fixation durations have not always been specified, but von Cranach's group found that the number of FG judgements given increased if 3 sec. rather than 1 sec. fixations were used (von Cranach, 1971*a*, *b*) and Cline (1967) used a much longer time of 10 sec. Amongst other variables, the number of trials for each fixation point may well be important, as even uncorrected practice may influence judgements where the task is a relatively unfamiliar one.

INTER-OBSERVER RELIABILITY OF GAZE DIRECTION JUDGEMENTS

A weaker measure of the validity of GD judgements is given by the amount of inter-observer agreement. Until recently, investigators have reported high values of this index in studies where GD was monitored during social interactions. The first systematic report comparing observers was by Exline (1963), who had pairs of observers judging the amounts of EG given by a sender to each of two receivers in three-person interactions. The judged EG totals were correlated across 10 triads, giving correlations of 0·97 and 0·98. Distance was unspecified, but the observers were at substantial angles to both sender–receiver axes. A more precise index of agreement was obtained by an 'act by act' comparison of the EG *v.* non-EG profiles as judged by the observers. This yielded only 81 per cent average agreement when discrepancies between records of 2 sec. or more were summed. Other studies have reported 'very high' or 'near perfect' agreement for judgement of gazing in dyads, with the observer positioned directly behind the receiver, and occasionally figures are given—Exline *et al.* (1965) report 88–98 per cent, Strongman & Champness (1968) report 95 per cent—but the criteria of agreement are not specified.

---

[*] Exline (1967; unpublished) indicated that in a study where gazes at 15 within-face targets *were* monitored with an eye camera, he found that confusion of targets by the judges was still common.

The above results are difficult to reconcile immediately with more recent figures. Mehrabian & Williams (1969) note that in a study where many behaviours were recorded, inter-observer reliability for EG judgements was only 0·50—one of their worst figures, but this study was based on video-tape records, and the observers were probably at a large angle to the sender–receiver axis. A more serious negative finding is due to Stephenson & Rutter (1970), who were doubtful of a result found by Argyle & Dean (1965), showing that amounts of mutual EG (or 'eye-contact') typically increased as the sender–receiver distance increased. Stephenson & Rutter showed that when the actual amount of EG given by a sender was kept constant, the greater the sender–receiver distance the greater the amounts of actual gazing at the ear or the shoulder that were judged by the observers to be EG. This controversy (see also Argyle, 1970; von Cranach *et al.*, 1968) does not concern us here except inasmuch as it raises methodological questions. Stephenson & Rutter's results agree with others (see Table 1) which indicate that gaze discrimination by observers gets poorer at greater distances, and is characterized by underestimation of deviation from the sender–receiver axis. The feature of their study which is pertinent to the present question is that they used pairs of observers and gave inter-observer correlations for each of 16 pairs across distances and fixation points. These ranged from −0·17 to +0·85, with a mean of only 0·43, and only half were significant correlations.

There is evidently some conflict between studies as to how well the observers can agree in judging GDs. Stephenson & Rutter (1970) suggest that the low reliabilities in their study may have been due to using naive observers. Other studies, by first training the observers may have induced shared, but not necessarily valid, judgement strategies giving higher agreement. This is consistent with the finding by White *et al.* (1970) that induced prior expectations could determine whether the observers recorded increase or decrease of EG with increasing distance. The evidence thus far does not appear to justify much confidence in judges' ability to discriminate between EGs and non-EGs very validly; but in all the studies so far discussed the experimental constraints were very high, senders' gazes either being programmed by the experimenter, or being free in artificial situations (e.g. with the receiver himself giving continuous EG to the sender). We must now consider whether validity may be higher when more naturalistic interactions are judged.

JUDGEMENT OF GAZE DIRECTION IN NATURAL VERSUS ARTIFICIAL CONDITIONS

Although the experimental evidence is conflicting, a number of experienced investigators have expressed high subjective certainty in discriminating EG from non-EG in situations where persons are interacting freely. The hypothesis which has been advanced to reconcile this with negative data is simply that in natural interactions we either give EGs or we look well away from the eyes of the receiver, i.e. we look at other targets in the head region rather rarely (Argyle, 1970; von

Cranach, 1971 *a*, *b*). If this *is* the case, then observers need only be able to discriminate FGs from non-FGs, which is relatively an easier task.

The hypothesis is difficult to test, as we cannot easily obtain objective records of interactors' GDs when these are not pre-programmed. In Munich consideration is now being given to the possible use of electro-oculographic techniques which might overcome the problem, but as yet there are no published studies bearing directly on the hypothesis. Indirect evidence can, however, be obtained by investigating inter- and intra-observer reliabilities in natural and pre-programmed visual interactions. Some unpublished data obtained by the present writer is relevant here, and will be presented briefly. Since its main function was as a personal check on reliability and validity of EG *v.* non-EG judgement in transcribing films (from a study still in progress), only the present writer and one other person acted as observers.

For the validity study two stooges were filmed on 16 mm black-and-white ciné film, with separate cameras facing each person and modified to film at 2 ft. per sec. They were operated separately: in the first run S_1 was filmed as he followed a prescribed pattern of 5 sec. fixations around the face of S_2, who was seated across a table from S_1 and giving a constant EG in return; in the second run roles were reversed and S_2 was filmed while following the same gazing sequence. Five fixation points (including an EG) were located within the other's face, and four were outside it (15 cm above, below, or to either side). Each fixation point was repeated eight times, with the seating postures of the senders varied on each occasion. The sender–receiver distance was between about 80 cm (both leaning forward) and 170 cm (both leaning back). One camera was fixed 270 cm from the face of S_1 at 25° to the sender–receiver axis, the other 210 cm from S_2 at 15° to the axis. The S_1 wore spectacles, and thus viewing conditions for the film of S_1 were in several respects poorer than for S_2. (This was reflected in the EG *v.* non-EG judgements, but only overall results based on both films together will be presented here.)

Both observers transcribed the films simultaneously but independently, using press buttons activating a pen-recorder. Two transcriptions of EG *v.* non-EG were made, with a four-day gap between viewings. A transcript of the actual fixation sequence was then made in conjunction with the original schedule of fixations. Inter-observer agreement was assessed by summing all discrepancies between records which exceeded 0·25 sec., over the duration of the films (approximately 6 min.). This was repeated using the 2 sec. criterion of Exline (1963) for comparison purposes. Agreement between the first transcriptions was 68·9 or 73·8 per cent, while between the second transcriptions it was 66·0 or 69·8 per cent. Agreement was thus relatively poorer than in Exline's study, even with his rather lax 2 sec. criterion; however, the second observer in the present study had no experience of judging GDs, and so no shared strategies had been induced by pre-training in this case.

Reliability of judgement between the two viewings was assessed using only the 0·25 sec. agreement criterion, reaching 76·4 per cent for the writer and 75·9 per cent for the other observer. Validity of judgements was assessed by comparing the judged with the actual transcripts, again for each viewing. Results were rather poor: the most successful discrimination was obtained by the writer on the first viewing, when 75·6 per cent of actual EG was judged to be EG, but so were 18·5 per cent of gazes not even directed at the receiver's face. In general, the results from the experimental films indicated that EG *v.* non-EG judgement was neither very reliable nor very valid. They suggest that when judgement conditions are less than optimal, and shared strategies are precluded, the discrimination is rather poor, although bearing in mind the fact that judgement was from filmed rather than live interactions, the inter-observer agreement was better than the average values found by Stephenson & Rutter (1970). This probably reflects the greater expertise of the observers in the present study, who were both trained in detecting fine movements, although only the writer was experienced in judging GD.

The same observers then made similar judgements of four 5-min. segments of research films, which had been made in the same physical set-up, but with naive subjects who were filmed without their prior knowledge during free conversations. The segments were selected randomly from the writer's collection, and were judged under similar conditions to before, once by both observers, and again four days later by the writer alone. These films were in all respects comparable to the experimental films except in that the senders were subject to no restraints in their gazing patterns. Results were appreciably better in that average inter-observer agreement was now 93·8 per cent (and the writer's self-reliability was 96·2 per cent). The evidence from this comparative study, although limited to only two observers, clearly tends to support the hypothesis that more accurate discrimination between EG and non-EG can be achieved when judging actual interactions than is generally possible in the controlled fixation situations.

There is no reason to suppose that actual discrimination ability of observers will vary between the two conditions cited, thus differences in accuracy are most likely to arise from the differences in gazing behaviour. The above data thus support the hypothesis that in real-life interactions we do tend either to give an EG or to look well away from the face of another person. If this is the case, then observer judgement can be expected to be fairly adequate in studies where EG *v.* non-EG is monitored, at least when visibility conditions are such that FG *v.* non-FG discrimination is good and the observers are experienced.

On the basis of the data which has been considered, we may conclude that we are not forced to disregard the findings of most studies where gazing during social interaction has been monitored, although particular situations are likely to lead to rather wide variations in accuracy. The importance of specifying the distance between interactors and the observation conditions has been shown to be considerable for the assessment of probable accuracy, and in common with other techniques relying on observer judgement there are risks of systematic bias. While training can yield high inter-observer reliability under suitable conditions, there is a need for new methods by which shared judgement strategies can be validated against external criteria if gazing variables are to be monitored accurately.

This paper is based on work carried out largely while the writer was a guest of Dr Mario von Cranach and the Max-Planck Society at the Institut für Psychiatrie in Munich. Mr J. H. Ellgring made helpful criticisms of a draft manuscript and provided information on unpublished studies. Dr J. D. Goss-Custard kindly served as an observer in the writer's own reliability/validity studies.

REFERENCES

ANSTIS, S. M., MAYHEW, J. W. & MORLEY, T. (1969). The perception of where a face or television 'portrait' is looking. *Am. J. Psychol.* **82**, 474–489.

ARGYLE, M. (1970). Eye-contact and distance: a reply to Stephenson and Rutter. *Br. J. Psychol.* **61**, 395–396.

ARGYLE, M. & DEAN, J. (1965). Eye-contact, distance, and affiliation. *Sociometry* **28**, 289–304,

ARGYLE, M. & KENDON, A. (1967). The experimental analysis of social performance. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, vol. 3. New York: Academic Press.

CLINE, M. G. (1967). The perception of where a person is looking. *Am. J. Psychol.* **80**, 41–50.

CRANACH, M. VON (1971*a*). Some factors in the perception of looking behaviour. *XIX Int. Congr. Psychol.* London: British Psychological Society.

CRANACH, M. VON (1971*b*). The role of orienting behaviour in human interaction. In A. H. Esser (ed.), *Environmental Space and Behaviour*. New York: Plenum Press.

CRANACH, M. VON & ELLGRING, J. H. (1971). The perception of looking behaviour. In M. von Cranach & I. Vine (eds.), *Expressive Movement and Nonverbal Communication*. London: Academic Press. (In press.)

CRANACH, M. VON, FRENZ, H. G. & FREY, S. (1968). Die 'angenehmste Entfernung' zur Betrachtung sozialer Objekte. *Psychol. Forsch.* **32**, 89–103.

DUNCAN, S., JR. (1969). Nonverbal communication. *Psychol. Bull.* **72**, 118–137.

ELLGRING, J. H. (1971). Die Beurteilung des Blickes auf Punkte innerhalb des Gesichts. *Z. exp. angew. Psychol.* (in press).

EXLINE, R. V. (1963). Explorations in the process of person perception: visual interaction in relation to competition, sex, and need for affiliation. *J. Personality* **31**, 1–20.

EXLINE, R. V., GRAY, D. & SCHUETTE, D. (1965). Visual behavior in a dyad as affected by interview content and sex of respondent. *J. Person. soc. Psychol.* **1**, 201–209.

GIBSON, J. J. & PICK, A. D. (1963). Perception of another person's looking behaviour. *Am. J. Psychol.* **76**, 386–394.

MEHRABIAN, A. & WILLIAMS, M. (1969). Nonverbal concomitants of perceived and intended persuasiveness. *J. Person. soc. Psychol.* **13**, 37–58.

STRONGMAN, K. T. & CHAMPNESS, B. G. (1968). Dominance hierarchies and conflict in eye-contact. *Acta psychol.* **28**, 376–386.

STEPHENSON, G. M. & RUTTER, D. R. (1970). Eye-contact, distance and affiliation: a re-evaluation. *Br. J. Psychol.* **61**, 385–393.

VINE, I. (1970). Communication by facial-visual signals. In J. H. Crook (ed.), *Social Behaviour in Birds and Mammals: Essays on the Social Ethology of Animals and Man*. London: Academic Press.

WHITE, J. H., HEGARTY, J. R. & BEASLEY, N. A. (1970). Eye contact and observer bias: a research note. *Br. J. Psychol.* **61**, 271–273.

Manuscript received 14 *July* 1970

Revised manuscript received 24 *November* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 332–341
Printed in Great Britain

Changes in Moral Belief among Sixth-form Boys and Girls over a Seven-Year Period in relation to Religious Belief, Age and Sex Difference

BY DEREK WRIGHT

Department of Psychology, University of Leicester

AND EDWIN COX

University of London Institute of Education

The results of replicating in 1970 a survey first conducted in 1963 show that there have been highly significant changes in the moral beliefs of sixth-form boys and girls in maintained grammar schools in England. The changes are greatest for sexual behaviour but extend to such issues as colour bar and gambling. In general, the change is away from the uncompromising condemnation of behaviour towards a more qualified, permissive or undecided position. Though moral and religious beliefs are strongly associated in both surveys, the results show that changes in moral belief are very largely independent of religious commitment. They were also found to be unrelated to age in this sample, though there were clear signs that girls had changed more than boys.

One of the most confident assumptions made today about English society is that moral standards are in decline. Yet adequate evidence is lacking. In order to demonstrate a change, it would be necessary to use the same sampling procedure and measuring instrument on separate occasions, and so far this does not appear to have been done.

A second assumption rivalling the first in its prevalence is that religious belief and practice provide the main defence against the erosion of moral standards, and that the decline in standards is functionally related to the decline in religion. In a recent survey (I.T.A., 1970) 89 per cent of a representative random sample of 1071 adults throughout Great Britain judged it true that 'Religion helps to maintain the standards and morals of society', and 69 per cent that 'To lead a good life it is necessary to have some religious belief'. A number of studies have consistently shown that people who express strong commitment to Christian belief and who attend church fairly frequently have more stringent and exacting moral beliefs than non-believers (see, for example, Gorer, 1955; Middleton & Putney, 1962; and, for a full discussion, Wright, 1971). These studies merely demonstrate an association between moral and religious belief, however; they do not show that there is any important functional link between them.

There are two main purposes behind the present paper. They are: to provide more reliable evidence than is usually available of the existence and nature of change in moral belief in one narrow but important segment of society; and to find out whether the extent of such change is in any way related to professed religious belief and practice. In 1963 the authors conducted a survey of the religious and moral beliefs of a large sample of sixth-form boys and girls from a random sample of 96 maintained grammar schools in England (see Wright & Cox, 1967*a*, *b*). Seven years later it was decided to repeat the survey in the same schools with the same indices of moral belief.

PROCEDURE

An attempt was made to contact the heads of all the schools used in the previous survey with the request that they select 25 second-year sixth-formers according to the method used before and give them a shortened and slightly modified version of the previous questionnaire under the same conditions as then. Anonymity was guaranteed.

The relevant questionnaire items

The original questionnaire contained a number of multiple-choice and open-ended questions on religious belief and practice, moral belief, and attitude towards religious education. The main alteration made to the questionnaire on the second occasion was to omit many of the open-ended questions. Changes in religious belief and practice and in attitude towards religious education will be reported elsewhere.

In the section devoted to moral judgement, subjects were presented with a list of forms of behaviour which are commonly made matters of moral evaluation and asked to rate them. The list contained the following: Gambling, Drunkenness, Smoking, Lying, Stealing, Premarital Sexual Intercourse, Suicide, Colour Bar, Taking Non-addictive Drugs (e.g. marijuana), and Taking Addictive Drugs (e.g. heroin). This list is identical to the one used in the 1963 survey except for the addition of the two items on drugs. The original list was based upon an informal survey of the moral issues most frequently discussed by boys and girls of the kind to be sampled, and in 1963 the taking of drugs was not one of them.

Subjects were asked to rate each item on the following scale: A, it is always wrong; B, it is usually wrong but is excusable in certain circumstances; C, it is usually excusable but is sometimes wrong; D, it is never wrong; E, I have not made up my mind. A space was provided below each item and subjects were urged to give reasons for the rating of their choice. Detailed analyses of these comments will be given elsewhere.

The indices of religious belief and practice to which reference will be made later in this paper include two five-point scales, one of degree of confidence in the existence of God and the other of degree of confidence that Christ was the Son of God, questions on the importance of the issue of life after death and belief in life after death, and finally a four-point scale of frequency of churchgoing and a three-point scale of frequency of private prayer.

The comparability of the two samples

It is possible to challenge the comparability of the two samples on at least three grounds. In the first place only 66 of the original 96 schools agreed to take part on the second occasion. Accordingly, boys and girls from these schools were compared with those from the remaining 30 on every questionnaire item on the 1963 data.* In no single instance did the difference begin to approach significance. This does not exclude the possibility that processes of change are operating differentially between the two groups of schools; but since this seems unlikely, it was concluded that though the sample of schools had been reduced in number, it nevertheless remained representative of the original population and would therefore allow generalization to it.

In the second place, 14 of the schools had been reorganized since the previous survey and the basis of entry to them changed. In the case of eight of them, the heads were confident that the change had not yet had any impact upon the composition of the sixth form. Three of the heads made no comment, and the remaining three simply reported that their sixth forms had become coeducational in the previous year or two. Since the numbers involved were relatively small and the extent of the influence of these changes uncertain, it was felt that reorganization did not introduce serious bias. Nevertheless, when taken together with the other ways in which schools can change—such as in size or in the areas from which they draw their pupils—it remains one of the main reasons for caution in interpreting the results.

---

* Chi-square analyses were used throughout.

Lastly, the first survey was conducted in October of the school year and the second in February–March. As a consequence the average age of the second sample is four to five months older than the first. In the 1963 sample no association could be found between age and moral judgement except that among girls there was a significant tendency for the oldest and the youngest to be more condemnatory of suicide. But the existence of an age difference between the samples makes it necessary to check for age relationships in the 1970 sample and for possible relationships between age and change in moral belief.

Subjects

These totalled 3850, of whom 2276 were in the 1963 survey and 1574 in the 1970 survey; 49·6 per cent of the 1963 sample and 48·1 per cent of the 1970 sample were boys, the remainder girls. Ages ranged from just under 17 years to just over 19.

RESULTS AND DISCUSSION

The main results from the 1963 and 1970 surveys are given in Table 1. In every instance and for both sexes there has been a significant shift in judgement, and in nearly every case the shift has been away from the unequivocal condemnation of behaviour towards a more qualified, lenient or undecided position. The single exception is smoking. Here the change has been towards a greater polarizing of the extremes, with, for each sex, an increased proportion judging it always wrong and an increased proportion judging it never wrong.

If we take the 'always wrong' category alone, then by far the greatest change in both sexes has occurred in relation to premarital sexual intercourse. Thereafter the rank orders for amount of reduction in this category are: *Boys*—drunkenness, colour bar, stealing, suicide, gambling, lying; *Girls*—drunkenness, colour bar, lying and stealing, gambling, suicide. A decline in those judging an action always wrong does not of course necessarily mean an increase in those judging it never wrong. In general, the increases in 'never wrong' are smaller than the reduction in 'always wrong'. As far as stealing and colour bar are concerned, there are appreciable increases in those judging them 'usually wrong', but negligible increases in those judging them 'never wrong'. If we consider increases in the 'never wrong' category alone, then once again the largest change for both sexes is over premarital sexual intercourse. Thereafter the rank orders for amount of increase are: *Boys*—drunkenness, suicide, gambling, smoking, colour bar, stealing and lying; *Girls*—suicide, smoking, drunkenness, stealing. Among boys the only sizable increases in the 'undecided' category were for suicide and colour bar, and among girls for these two and premarital sexual intercourse and drunkenness.

In view of these complexities the most plausible interpretation to put on the results would appear to be not that there has been a shift away from morality so much as an increased awareness of the complexity of moral problems and a disinclination to adopt extreme and uncompromising positions. And it remains true that stealing, colour bar and—we may now add—the taking of addictive drugs receive strong moral condemnation. The most notable exception to this is premarital sexual intercourse, where the change is not only massive but appears to be in the direction of removing the issue from moral consideration completely, especially among boys. However, to understand the nature of the change on this item it is necessary to take account of the comments made by the subjects. A surprisingly high proportion of

Table 1. *The percentages of the total numbers of boys and girls endorsing the different rating categories in the 1963 and 1970 samples*

	Always wrong	Usually wrong	Sometimes wrong	Never wrong	Undecided
BOYS					
Gambling					
1963	10·7	16·9	57·0	8·3	7·1
1970	8·8	16·4	55·7	11·5	7·5 (0·01)
Drunkenness					
1963	30·6	46·5	17·9	2·0	3·0
1970	17·7	35·7	34·5	8·3	3·8 (0·001)
Smoking					
1963	14·9	10·2	32·4	31·6	10·9
1970	20·9	11·4	21·7	34·5	11·6 (0·001)
Lying					
1963	19·1	73·9	5·2	0·5	1·3
1970	18·4	70·2	6·9	1·7	2·9 (0·001)
Stealing					
1963	70·8	27·5	0·6	0·3	0·8
1970	65·0	30·5	1·6	1·6	1·3 (0·001)
Premarital sexual intercourse					
1963	28·6	27·6	20·5	10·2	13·1
1970	10·3	12·8	30·4	33·7	12·8 (0·001)
Suicide					
1963	32·0	28·7	13·0	15·2	11·1
1970	28·5	23·6	10·6	19·3	18·0 (0·001)
Colour bar					
1963	74·0	16·8	3·2	2·4	3·6
1970	65·6	18·1	3·7	4·9	7·7 (0·001)
Non-addictive drugs					
1963	—	—	—	—	—
1970	29·3	17·8	18·9	19·8	14·1
Addictive drugs					
1963	—	—	—	—	—
1970	65·9	14·0	4·1	7·8	8·2
GIRLS					
Gambling					
1963	19·4	27·1	37·9	3·5	12·1
1970	13·2	24·9	46·5	2·9	12·5 (0·001)
Drunkenness					
1963	44·9	44·4	7·8	1·1	1·8
1970	22·4	47·4	24·0	2·0	4·2 (0·001)
Smoking					
1963	12·3	8·9	36·2	24·4	18·2
1970	18·2	11·5	26·3	30·2	13·7 (0·001)
Lying					
1963	23·8	72·8	2·2	0·3	0·9
1970	17·3	74·5	6·2	0·1	1·8 (0·001)
Stealing					
1963	71·2	27·8	0·5	0·1	0·4
1970	64·7	32·8	0·7	0·2	1·5 (0·01)

Table 1—*continued*

	Always wrong	Usually wrong	Sometimes wrong	Never wrong	Undecided
Premarital sexual intercourse					
1963	55·8	25·2	6·6	2·4	10·0
1970	14·6	19·5	30·0	17·7	18·2 (0·001)
Suicide					
1963	33·4	32·7	10·1	8·7	15·1
1970	29·4	26·1	10·2	17·2	17·1 (0·001)
Colour bar					
1963	85·3	8·9	1·6	1·5	2·7
1970	76·4	14·7	1·7	1·5	5·7 (0·001)
Non-addictive drugs					
1963	—	—	—	—	—
1970	41·6	19·6	17·3	7·8	13·6
Addictive drugs					
1963	—	—	—	—	—
1970	73·2	14·2	2·6	2·7	7·3

those who judged premarital sexual intercourse never wrong, nevertheless added that this was only so when the couple were in love or had a sense of responsibility towards each other. The 1970 sample seem to be as insistent as the 1963 sample in condemning 'sex for kicks' and indiscriminate promiscuity. It seems plausible to argue that the change in rating represents a decided weakening in the link between sexual intercourse and marriage in the minds of the subjects, while the comments witness if anything to a strengthening of the link between sex and affection.

Sex differences

In the 1963 survey girls were found to be significantly more severe in their judgements on all issues except stealing, where the difference was in the same direction but not significant, and smoking, where the difference was significant but in the opposite direction. The pattern of association was closely similar to this in the 1970 survey. Girls were significantly stricter in their judgements in regard to gambling, drunkenness, premarital sexual intercourse, colour bar, and the taking of addictive and non-addictive drugs. For lying there was a significant tendency ($P < 0·05$) for boys to occupy the more extreme positions and a similar tendency for smoking that was not significant ($P < 0·10$). There was no sign of any significant association between sex difference and assessment of suicide and stealing.

As we saw earlier, the rank orders for amount of decline in the proportions who endorse the 'always wrong' category were similar for the sexes. In every instance, however, the decline was greater in amount for girls than for boys, and in the case of premarital sexual intercourse it was more than twice as great. On the other hand, on only two of the issues did girls show a greater increase in the 'never wrong' category. As a crude index of total amount of change, the percentage differences between the 1963 and 1970 samples were summated over all rating categories for each item and for each sex. On this index, girls showed more change on all items except drunkenness and suicide. A Wilcoxon T test just failed to reach significance at the 5 per cent level, two-tailed. We can conclude that there is some general trend for girls to change more than boys, but that this change cannot be simply interpreted as a movement towards greater similarity to the boys.

Age and subject studied at school

As with the 1963 sample, one moral issue was related to age in the 1970 sample, only this time it was stealing. There was a tendency for younger boys to be more severe in their condemnation of it. The general lack of association with age in both samples makes it unlikely that one age group will have changed more than another. Comparisons were nevertheless made between the two samples on every item for each three-month age grouping. Close scrutiny of the resulting pattern of associations failed to reveal any sign of a tendency for one age group to change more than any other with the single exception of smoking among boys. Here the pattern was for the youngest groups to show a significant shift towards greater condemnation, the intermediate groups no significant difference, and the oldest group a significant change towards greater leniency. It is possible that this reflects a tendency for older boys to be more often addicted to smoking.

In the 1963 sample no association could be found between subject studied at school and any of the moral judgement items. Among boys in the 1970 sample there were weak trends for arts pupils to be more condemnatory of gambling ($P < 0{\cdot}05$) and lying ($P < 0{\cdot}10$); among girls there was a significant tendency for arts students to be more severe on gambling ($P < 0{\cdot}01$) and science students on smoking ($P < 0{\cdot}01$). There was also a weak, non-significant tendency for scientists to be more condemnatory of colour bar.

Religious belief and practice

In the 1963 survey strong and consistent associations were found between religious belief and practice and severity of moral judgement. The only exceptions were that among boys belief in God and smoking were unrelated and among girls there was no association between colour bar and belief in God, churchgoing, and private prayer. Table 2 gives the significance levels of the corresponding tests for the 1970 sample.

It is clear from Table 2 that though the association between religion and moral judgement is still strong, it has been weakening somewhat over the last seven years. Among boys there is little, if any, connexion between religious belief and practice and judgements on smoking, lying, stealing and colour bar, and among girls between religion and smoking and colour bar. It is interesting to find that colour bar, which was least strongly related to religion in the 1963 sample, is even less related in the later sample. It is worth adding that in the United States there is strong evidence of a positive association between racial prejudice and conventional religiosity (e.g. Allport & Ross, 1967).

As we pointed out earlier, religious belief and practice are frequently said to be the necessary foundation of morality without which moral standards are sure to crumble. It is not unreasonable to argue that if, in relation to the present study, it were found that the proportions of believers who endorse the stricter moral rating categories have not declined whereas the proportions of disbelievers have, then the plausibility of this assertion would be enhanced. Conversely, if religious believers have changed as much as, or more than, disbelievers, then its plausibility would be diminished. It should be added, however, that the value of the test is somewhat weakened by the fact that, to be valid, it requires that the content of the various religious ideologies in relation to the moral issues should not have changed over the last seven years, and this cannot easily be checked.

Table 2. *Significance levels of chi-square tests relating indices of religious belief and practice to moral judgement in the* 1970 *sample*

(In general, when the results are significant, religious subjects tend to be more strict in their moral judgement. Exceptions are marked with an asterisk.)

	Gambling		Drunkenness		Smoking		Lying		Stealing	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Belief in God	0·001	0·001	0·001	0·001	n.s.	n.s.	n.s.	0·001	n.s.	0·001
Belief in Christ	0·001	0·001	0·001	0·001	n.s.	n.s.	n.s.	0·001	0·01	0·05
Importance of life after death	0·001	0·01	0·001	0·01	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Belief in life after death	0·05	0·02	0·001	0·05	0·05	n.s.	n.s.	n.s.	n.s.	n.s.
Churchgoing	0·001	0·001	0·001	0·001	n.s.	n.s.	n.s.	0·01	0·05	0·01
Private prayer	0·001	0·001	0·001	0·001	n.s.	0·05	n.s.	0·02	n.s.	0·05

	Premarital sexual intercourse		Suicide		Colour bar		Taking non-addictive drugs		Taking addictive drugs	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Belief in God	0·001	0·001	0·001	0·001	(0·10)	n.s.	0·001	0·001	n.s.	0·01
Belief in Christ	0·001	0·001	0·001	0·001	n.s.	n.s.	0·001	0·001	0·001	0·001
Importance of life after death	0·001	0·001	0·001	0·001	0·05	0·01*	n.s.	0·001	n.s.	n.s.
Belief in life after death	0·001	0·001	0·02	0·05	0·05	n.s.	n.s.	(0·10)	n.s.	n.s.
Churchgoing	0·001	0·001	0·001	0·001	n.s.	0·05*	0·01	0·001	0·05	0·01
Private prayer	0·001	0·001	0·01	0·001	n.s.	n.s.	0·001	0·001	n.s.	0·001

In testing the claim, then, the two samples were compared in turn on each of five degrees of belief in God and in Christ, four frequencies of reported churchgoing, and three frequencies of reported private prayer. In making these comparisons the sexes were combined. The results are given in Table 3; they do not present a simple picture. Each moral issue will therefore be taken separately.

Gambling. It is difficult to discern any consistent tendency in these results for believers or non-believers to have changed more.

Drunkenness. At all levels of belief and practice, the changes were highly significant ($P < 0.001$).

Table 3. *Significance levels of change from the 1963 to the 1970 sample by level of religious belief and practice, sexes combined*

	Gambling	Drunkenness	Smoking	Lying	Stealing	Premarital sexual intercourse	Suicide	Colour bar
Belief in God								
1. Confident belief	n.s.	0·001	0·001	0·02	0·01	0·001	0·01	0·01
2. Fairly confident belief	0·001	0·001	0·001	0·001	n.s.	0·001	0·001	0·001
3. Undecided	n.s.	0·001	0·001	n.s.	n.s.	0·001	0·001	0·001
4. Fairly confident disbelief	0·05	0·001	0·02	n.s.	n.s.	0·001	0·001	n.s.
5. Confident disbelief	n.s.	0·01	n.s.	n.s.	n.s.	0·001	0·001	0·05
Belief in Christ								
1. Confident belief	n.s.	0·001	0·001	0·02	0·01	0·001	0·05	0·001
2. Fairly confident belief	n.s.	0·001	0·001	0·01	n.s.	0·001	0·05	0·001
3. Undecided	n.s.	0·001	0·001	n.s.	0·02	0·001	0·01	0·001
4. Fairly confident disbelief	n.s.	0·001	n.s.	n.s.	n.s.	0·001	0·01	n.s.
5. Confident disbelief	n.s.	0·001	0·01	n.s.	n.s.	0·001	0·001	0·05
Churchgoing								
1. Weekly	n.s.	0·001	0·001	0·01	n.s.	0·001	n.s.	0·001
2. Monthly	n.s.	0·001	n.s.	0·01	n.s.	0·001	0·001	0·001
3. Sometimes	n.s.	0·001	0·001	0·01	n.s.	0·001	0·001	0·01
4. Never	0·05	0·001	0·001	n.s.	n.s.	0·001	0·001	0·05
Private prayer								
1. Daily	0·05	0·001	0·001	0·001	0·05	0·001	0·05	0·001
2. Sometimes	n.s.	0·001	0·001	0·001	0·001	0·001	0·001	0·001
3. Never	0·01	0·001	0·001	n.s.	n.s.	0·001	0·001	0·01

Smoking. This was the only item in which the direction of change was partly towards stronger rejection. There is perhaps a faint tendency for less change among the non-believers.

Lying. The pattern is consistent. There are highly significant changes in the direction of a more permissive point of view among those who strongly believe and attend church and pray frequently. Among non-believers there has been virtually no change for any of the tests.

Stealing. The only significant shifts are to be found among the more religious subjects.

Premarital sexual intercourse. At all levels of belief and practice the change was highly significant ($P < 0{\cdot}001$), and there was no indication that any one level of religious belief and practice had changed more than any other.

Suicide. Nearly all tests were significant, but there were traces of somewhat less significant change among believers.

Colour bar. Most tests yielded highly significant results. There were, however, fairly clear signs of less significant change among the non-religious.

To sum up, there are some barely discernible signs that religious subjects have changed less than the non-religious over suicide, and that they have increased more in strictness over smoking; on the other hand they have changed as much as the non-religious in regard to gambling, drunkenness and premarital sexual intercourse and more than the non-religious in regard to lying, stealing and colour bar. It can hardly be said that these results offer any support for the claim that religious commitment conserves moral belief. On the contrary, since it is unlikely that religious ideologies have changed in recent years in regard to lying, stealing and the colour bar, it can be argued that the claim has been clearly disconfirmed. It is true that the nature of the evidence is such as to render any interpretation tentative and speculative, and the issues raised are too complex to be dealt with adequately here. Nevertheless, it is plain that though strictness of moral belief is associated with religious commitment in both samples, *changes* towards a more permissive point of view are in no way delayed or impeded by such commitment. It is notable that of all moral issues premarital sexual intercourse has been found among those most strongly related to religious belief in this and other surveys, yet it is the one on which change has been greatest, and as much among believers as non-believers. Though modifications in the content of religious beliefs in relation to sexual behaviour may have contributed to this change, it seems unlikely that they are the main cause; indeed the authorities within the various denominations appear to have modified their position very little on this issue. The implication of all this is that, though factually associated, religious and moral beliefs are not functionally related.

REFERENCES

ALLPORT, G. W. & ROSS, J. M. (1967). Personal religious orientation and prejudice. *J. Person. soc. Psychol.* 5, 432–443.

GORER, G. (1955). *Exploring English Character*. London: Cresset Press.

I.T.A. (1970). *Religion in Britain and Northern Ireland*. London: Independent Television Authority.

Middleton, R. & Putney, S. (1962). Religion, normative standards, and behavior. *Sociometry* **25**, 141–152.
Wright, D. (1971). *The Psychology of Moral Behavior*. Harmondsworth: Penguin Books.
Wright, D. & Cox, E. (1967*a*). Religious belief and coeducation in a sample of sixth form boys and girls. *Br. J. soc. clin. Psychol.* **6**, 23–31.
Wright, D. & Cox, E. (1967*b*). A study of the relationship between moral judgment and religious belief in a sample of English adolescents. *J. soc. Psychol.* **72**, 135–144.

Manuscript received 19 *August* 1970

Revised manuscript received 15 *March* 1971

Br. J. soc. clin. Psychol. (1971), **10**, pp. 342–345
Printed in Great Britain

The Effect of Mood on Performance on the Modified New Word Learning Test (Walton–Black)

BY JOHN D. TEASDALE AND J. GRAHAM BEAUMONT*

Psychology Department, Institute of Psychiatry, University of London,
and Drug Dependence Clinical Research and Treatment Unit,
Maudsley and Bethlem Royal Hospitals, London

The Modified New Word Learning Test (MNWLT) is a psychological test of brain damage which has 'appeared capable of differentiating organics with general cortical damage from functionals and normals with a negligible degree of misclassification' (Walton *et al.*, 1959). In relation to differential diagnosis between organics with general cortical damage and depressives, Walton *et al.* (1959) concluded that while 'depressives often present a major problem in differential diagnosis . . . no score below 6 was obtained' (scores of 5 and below are regarded as 'organic' providing that the patient is not ESN or mentally defective), 'though the results of both the neurotic and psychotic depressives were generally lower than those of other non-brain-damaged psychiatric patients'. The mean for the 16 neurotic depressives in their study was 8·5 (range 6–10).

The present study reports results from two patients, which suggest that the effect of depressed/anxious mood on MNWLT scores may lead to misclassification of patients as organic.

Patient 1 was a 24-year-old woman admitted for treatment of barbiturate dependence of 20 months' duration. She was referred for psychological investigations of brain damage.

She was first tested during the afternoon, when she obtained a WAIS Full-Scale IQ of 104, with a normal subtest pattern, and a score of less than 4 on the MNWLT, testing being discontinued as the patient became agitated at her poor performance, and a clearly 'organic' score had already been obtained.

The patient did not appear depressed during testing, but later reported that she had felt depressed, and that the difficulties of memory, reading and problem-solving of which she had complained were always worse in the afternoons.

Accordingly, the MNWLT was administered on two mornings, once in the prescribed manner, and once using nonsense words invented by Walton (unpublished). She obtained scores of 8 and 9, clearly in the 'non-organic' range, and said she felt less depressed than at previous testing. It was hypothesized that the patient was in fact not organically impaired, but that she showed variation in MNWLT scores related to diurnal variation in mood.

To test this hypothesis, the patient was administered first the 'right now' version of the Multiple Affect Adjective Check List (MAACL) (Zuckerman *et al.*, 1964) and then the MNWLT on each of three consecutive mornings and afternoons in each of two consecutive weeks. The words for the MNWLT were selected from the Thorndike & Lorge (1944) word count to be just beyond the vocabulary attainment of the patient.

---

* Present address: Psychology Department, University College, Cardiff.

The patient received antidepressant medication (imipramine) from six days before the beginning of the experiment. Imipramine is thought to be effective from 10 to 20 days after first administration. Thus any drug effect was unlikely to confuse the first series of trials, but might have had some effect on the second series.

The results are shown in Table 1. They show a powerful inverse relationship between MNWLT scores and MAACL Depression, Anxiety and Hostility scores for trials 1–6 (first week), but no obvious relationship for trials 7–11 (second week). The correlation between MAACL Depression and MNWLT scores was −0·89 (d.f. = 4; $P < 0{\cdot}05$) in the first week, but −0·18 (d.f. = 3; n.s.) in the second week. It is suggested that this difference between the two weeks reflects the confounding effect of imipramine, which probably did not become active until the second week. Trial 12, the last scheduled trial, was omitted as the patient made a determined suicide bid after trial 11.

Table 1. *Patient 1: covariation of MNWLT and MAACL scores*

			MAACL		
Day	Trial	MNWLT	Depression	Anxiety	Hostility
1 (p.m.)	1	6	16	6	8
2 (a.m.)	2	9	7	1	4
2 (p.m.)	3	6	14	11	9
3 (a.m.)	4	10	9	2	6
3 (p.m.)	5	6	17	9	13
4 (a.m.)	6	8	9	5	4
8 (p.m.)	7	9	16	10	9
9 (a.m.)	8	6	18	10	11
9 (p.m.)	9	7	21	11	13
10 (a.m.)	10	8	11	3	4
10 (p.m.)	11	6	13	5	6

Results from the first week are consistent with the hypothesis that in this patient MNWLT socres were related to mood, and that the original 'organic' score resulted from depressed/anxious mood rather than brain damage. Both mood and MNWLT scores could, of course, be independently determined by diurnal variations in some third variable.

Patient 2 was a 39-year-old man admitted for treatment of barbiturate dependence of 20 years' duration. He was referred for psychological investigations of brain damage, having complained of poor memory and concentration. His WAIS Full-Scale IQ was 94, Verbal IQ 117, and Performance IQ 83, the Verbal–Performance discrepancy being abnormal at the 0·1 per cent level. On the first administration of the MNWLT, a score of less than 3, falling clearly in the 'organic' range, was obtained. At this time, the Psychiatric Out-Patient Mood Scale (POMS) (McNair & Lorr, 1964), completed prior to MNWLT administration, indicated high levels of self-reported tension, depression and fatigue (Table 2, occasion 1). The patient was retested 26 days later, as clinical observation had suggested a possible improvement in mental state. The MNWLT score was less than 5, still in the 'organic' range, and the POMS scores obtained prior to MNWLT administration in fact suggested little change in mental state (Table 2, occasion 2).

Table 2. *Patient 2: scores on MNWLT and POMS on four occasions*

Occasion	1	2	3	4
MNWLT	<3	<5	7	8
POMS				
Tension	27	27	5	18
Depression	25	25	3	12
Confusion	12	7	1	6
Anger	7	8	4	4
Fatigue	20	16	2	10
Vigour	6	8	17	13
Friendliness	15	14	15	15

Sixteen days later, the tests were readministered 2 hr. after the administration of 200 mg. pentobarbitone. The results (Table 2, occasion 3) suggested that on this occasion the patient's mood was much improved, and the MNWLT score of 7 was in the 'non-organic' range. In an attempt to show that this improvement in MNWLT score was the result of barbiturate-induced improvement in mood rather than a practice effect, both tests were readministered the following day in the non-drug condition. However, mood was still good (Table 2, occasion 4), as the patient was pleased at his imminent discharge and resumption of regular barbiturate consumption, and the MNWLT score of 8 was still in the 'non-organic' range. Different sets of words of equivalent difficulty were used for the MNWLT on each occasion.

Two hypotheses are consistent with the data: (1) the patient was brain-damaged and improvement in MNWLT scores was the result of 'learning to learn'; (2) the patient was not brain-damaged, but originally obtained 'organic' MNWLT scores as a result of his depressed/anxious mental state. The second of these can also accommodate the results from the first patient, and it is suggested that this is the hypothesis that can most parsimoniously explain the results from the two patients put together. Both patients were drug addicts, but the findings suggest a need to re-examine the use of the MNWLT with all patients who might be depressed. They may also have implications concerning the nature of depression.

The authors would like to thank Dr P. H. Connell, Director of the Drug Dependence Clinical Research and Treatment Unit, for encouragement and permission to study the patients in his care. J. G. Beaumont was supported by an award for Further Education in Medical Sciences from the Medical Research Council when this study was conducted.

REFERENCES

McNair, D. M. & Lorr, M. (1964). An analysis of mood in neurotics. *J. abnorm. soc. Psychol.* **69**, 620–627.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of* 30,000 *Words*. New York: Teachers College, Columbia University Press.

WALTON, D., WHITE, J. G., BLACK, D. A. & YOUNG, A. J. (1959). The Modified Word-Learning Test: a cross-validation study. *Br. J. med. Psychol.* **32**, 213–220.

ZUCKERMAN, M., LUBIN, B., VOGEL, L. & VALERIUS, E. (1964). Measurement of experimentally induced affects. *J. consult. Psychol.* **28**, 418–425.

Manuscript received 4 *December* 1969

Revised manuscript received 17 *March* 1971

Br. J. soc. clin. Psychol. (1971), 10, pp. 346–354
With 1 text-figure
Printed in Great Britain

Inter- and Intra-modal Transfer in Children with Modality Specific and General Handicaps

BY N. O'CONNOR AND B. HERMELIN

Developmental Psychology Unit, Medical Research Council,
Drayton House, Gordon Street, London

Blind, normal, subnormal and deaf children were taught to discriminate between touch stimuli lasting either 6 sec. or 2 sec. Appropriate motor responses to the long and the short stimulus were taught to each subject. When the basic touch task had been learned, the subjects were asked to do a transfer task. In this they also had to distinguish between stimuli of 2 sec. or 6 sec. duration. The transfer stimuli were two touch signals of a different kind from that in the learning task, air-puffs, and sounds or lights for the blind and deaf and their control groups respectively. The results showed that the initial touch discrimination transferred only to another kind of touch stimulus, but not to light or sound. There was also no transfer to the air-puff, except in the deaf children. Although most subjects could verbalize the solution of the learning task, they could not transfer it to other modalities. Absence of speech was no special handicap, as the deaf did at least as well as other groups.

Experiments on crossmodal integration or crossmodal transfer have been bedevilled by a confusion of terminology. By crossmodal transfer one may mean at least three different things. One of these is crossmodal matching. In this procedure the subject is allowed to explore an object through one sense, e.g. touch, and then subsequently to identify it from among other objects when presented in a different modality such as vision. A second is crossmodal discrimination. Two stimuli are presented and discriminated in one modality, and immediately following they are presented for discrimination in another. The effect of one discrimination task on the other is measured. A third type of crossmodal transfer involves the transfer of a principle. In this case a variety of procedures can be used, but essentially a discrimination based on a dimension or cue in one modality is transferred to a different modality using the same dimension or cue but in relation to different discriminanda. Previous studies by the authors investigated crossmodal matching or discrimination, while this study is concerned with the transfer of a principle within and across modalities.

The nature of the physiological basis for crossmodal integration has been of concern to experimenters in studies with animals and man. Geschwind (1965) has proposed that the inferior parietal lobule acts as an intermodal association area. This area is relatively developed in man, although slow to myelinate. It does not receive most of its impulses from the primary projection areas of touch, vision and hearing, but from the adjacent sensory association areas. In this sense, this parietal area might be the anatomical basis for the integration of information from all modalities. Geschwind argues against the view that speech is the underlying basis of crossmodal integration, suggesting that the development of speech rather depends on the existence in the human of a capacity to integrate visual, auditory and motor stimuli.

A biologically determined view of crossmodal integration seems to be confirmed by the results of psychological experiments, which generally establish the existence of crossmodal integration in adult humans, but not in young children or in animals. Ettlinger (1967), for example, failed to find crossmodal transfer in monkeys.

As crossmodal transfer is thought to develop with age in humans (Birch & Lefford, 1963), it has sometimes been supposed that verbal mediation which develops concurrently, is necessary for crossmodal coding to occur. This view seems to have been generated in part by Russian work such as that of Merlin (1957), on the transfer of conditioned reflexes from one signal system to the other, or that of Luria (1961) in which verbal mediation is seen as facilitating coding. However, there would appear to be no conclusive evidence of the value of verbal mediation in transfer of this kind for young children. When O'Connor & Hermelin (1961) presented simple line drawings of common objects to subnormals, the names of these objects were subsequently better recognized than were the same line drawings if presented for recognition along with others. However, this case is one in which, as in the Russian work, transfer was from vision to speech itself. In a subsequent experiment (Hermelin & O'Connor, 1964) discrimination of shapes was tested in vision after discrimination by touch and *vice versa*. We concluded: 'As the groups differed in verbal ability but not in their capacity to transfer from one modality to another, a hypothesis attributing such crossmodal effects to verbal coding must be rejected.' Thus the processes involved in crossmodal transfer of a learned principle of solution remain an open issue. If crossmodal transfer occurs, no clear evidence is available to indicate whether or not this transfer has an anatomical, mediational or a developmental foundation.

Evidence from experiments with children is inconclusive. Smith & Tunick (1969) carried out an experiment with retarded children, where discrimination had to be transferred from touch to vision and vision to touch. When the same cues or stimulus objects were used, transfer was achieved across modalities. No transfer was found, however, when the discrimination was to be transferred using the same dimension of discrimination (e.g. shape), but different objects. Blank & Bridger (1966) reported positive transfer of discrimination between sounds and lights provided that there was adequate verbalization during training. Blank, Altman & Bridger (1968) subsequently demonstrated positive transfer between touch and vision even in children unable to name the stimuli. While Birch & Belmont (1965) have reported deficits in matching auditory with visual patterns in cerebral palsied children, Rudel & Teuber (1968) found brain-damaged and retarded children not to be different from controls in visual–tactile and tactile–visual shape recognition.

Thus the relevance of developmental factors, brain damage and verbal mediation remains in question. It seemed appropriate to consider crossmodal integration in terms of developmental deficits associated with diffuse brain damage, and also in relation to such speech impairment as might arise as a result of deafness. We therefore selected deaf, subnormal and normal children matched for mental age. We also included blind children to investigate the effects of a specific sensory handicap on crossmodal transfer. According to Piaget, the first symbolic function develops before speech on the basis of visual imagery. Without this foundation in imagery, concept development may be impaired. Presumably speech in these circumstances would remain a relatively independent entity not closely associated with thought.

METHOD

As all children in the present experiment had adequate tactile sensitivity, touch was used as the modality for the initial discrimination learning. There were three stimuli in the transfer tasks, pressure on the palm delivered by a softer instrument than that used in learning, air-puffs to the palm and either sounds or lights. As a temporal dimension is common to all these modalities, duration was chosen as the relevant cue for discrimination, and a short stimulus had in all instances to be distinguished from a longer one. Thus the aim of this experiment was to study the effect of different handicaps on different kinds of transfer. The deficiencies involved were blindness, deafness and mental subnormality. The transfer tasks were of three kinds: transfer along a scale, transfer between dimensions and transfer across modalities. A scale was defined as the range between two stimulus points, in our case from soft to hard. A dimension was regarded as any perceivable quality of stimuli within any one sensory modality. For example, colour and form are both dimensions within the visual modality. A modality was assumed to be dependent on a particular sense organ, such as sight, touch or hearing.

A learned time discrimination was thus to be transferred in three ways. The scalar transfer was from a hard tactile stimulus to a soft one. The dimensional transfer was from one kind of touch to another (air-puff), and the crossmodal transfer was from touch to light or from touch to sound.

Stimulus equivalence was checked by presenting these temporal discriminations as learning tasks to separate matched groups.

Subjects

Sixty children acted as subjects. These were 10 deaf and 10 blind children, and 20 normals as well as 20 subnormals. None of the blind children had more vision than 3/60 and the deaf children had a minimum hearing loss of at least 60 db in each ear in the range of 200 to 2000 Hz. The chronological ages of the other groups were matched with the mental ages of the subnormals. These ages ranged from 6:9 years to 8:3 years, with an overall mean of 7:4 years. Analysis of variance showed that the chronological ages of the other groups and the mental ages of the subnormals were statistically identical. The chronological ages of the subnormals ranged from 8 years to 12 years, with a mean of 10 years 3 months. All children attended schools and lived at home.

In order to obtain 60 children for the transfer task, 128 had to be tested on the initial learning problem. Table 1 shows the number of children in each group who did not reach learning criteria, and were therefore excluded.

Table 1. *Rejection rates for initial learning task*

Groups	Accept	Reject
Blind	10	5
Deaf	10	13
Subnormal controls (for blind group)	10	18
Subnormal controls (for deaf group)	10	18
Normal controls (for blind group)	10	7
Normal controls (for deaf group)	10	7
Total	60	68

Chi-square showed that the blind had the lowest rejection rate. This was significantly different from that of the deaf ($P < 0{\cdot}05$) and the subnormal ($P < 0{\cdot}02$), but not from that of the normal. The normal had the next lowest rejection rate and differed from the subnormal ($P < 0{\cdot}02$). The deaf and the subnormal had the highest rejection rates. Those rejected did not differ statistically in chronological or mental age from those chosen to go on to the transfer tasks.

Apparatus

The apparatus for delivering the touch, air, light and sound stimuli is presented in Fig. 1. It consisted of a structure 67·5 cm long by 45·0 cm wide. At the experimenter's side it was 30·0 cm high and sloped down to 15·0 cm at the subject's side. There were three compartments, each 22·5 cm wide, which trisected the structure lengthwise. The subject put his left hand into a guide in the left-hand aperture and his right into a foam-rubber upholstered compartment on the right side.

Touch and air stimuli were delivered to the palm of the left hand, and the right hand made the response of knocking or rubbing in the right-hand compartment. The rubber padding in this compartment reduced the sound and the cover prevented the subject from seeing the hand.

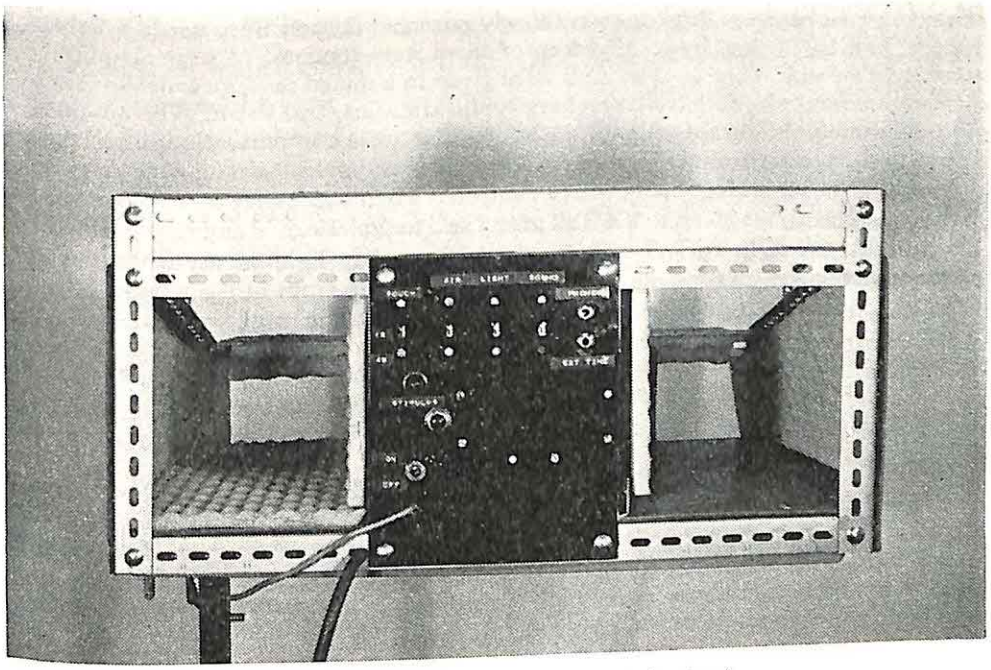


Fig. 1. Apparatus. (experimenter's view).

The centre compartment contained the electronic timer, tone generator, motor for rotating the tactile probe and air-puff relay. The touch stimulus apparatus consisted of a mains operated induction motor operating with reduced torque by means of a series resistor. The probe assembly was brought into contact with the subject's hand by moving the motor assembly forward by means of a solenoid. The motor and solenoid were both operated by the timer. Air (nitrogen) was delivered from a cylinder at a pressure of 2 lb/sq. in. by means of a regulator valve, and timed on and off by a tyre valve operated by a relay. The sound stimulus was a 400 Hz tone delivered via headphones to the subject at a level of 70 db. A mains strip light mounted along the top of the apparatus provided the light stimulus, this being matched in intensity to 70 db by means of a series resistor according to Stevens & Galanter (1957). The experimenter faced a control panel on which he could select a short (2 sec.) or long (6 sec.) presentation for each of the four stimuli.

Procedure

Subjects in each group learned a task which involved the discrimination of two separate durations (2 sec. and 6 sec.), presented as separate and successive tactile stimuli to the palm of the left hand. The subject sat with his right and left hands in the openings in the apparatus.

He was first taught two responses. These were two easily learned movements of the right hand, knocking and rubbing. Each subject first demonstrated that he could do both. He was then told that he would be touched on the left hand and that there would be two kinds of touches. These were then demonstrated to him. He was then told: 'These touches tell you when to knock and when to rub; you must find out which to do'. The touches were administered to the middle of the palm of the subject's left hand by a silently operating electric motor which turned a smooth-ended rod and brought it into contact with the palm. A rotary stimulus was used to avoid too rapid accommodation to a static touch stimulus. Whichever response the child made to the first stimulus was accepted as correct, and thereafter consistency of response was rewarded by the experimenter saying 'good', while incorrect responding was discouraged by saying 'no'. Training proceeded to a criterion of 51 correct responses in 54 consecutive trails. If criterion had not been reached by 90 trials, the subject was dropped from the experiment. The stimuli were administered in a randomly ordered series of 18 signals, nine long (6 sec.) and nine short (2 sec.). The series was repeated as often as necessary.

When criterion had been reached the subjects proceeded immediately to the transfer tasks. They were told that they would now play a similar game and they simply had to continue to knock and rub as before. Fifty-four randomly arranged stimuli were used. Of these, half were 2 sec. and half 6 sec. long. Eighteen of these were touches, 18 were air-puffs and 18 sound or light stimuli. These 54 stimuli were given in a mixed random order so that a short touch stimulus might be followed by a long sound stimulus, and so on. The air-puffs were given to the same point on the palm of the left hand as were the touch stimuli, and the sound stimuli were delivered through earphones. Touch, air and sound stimuli were given to blind children and their normal and subnormal controls. The other series consisted of touch, air and light stimuli and was given to the deaf group and to their normal and subnormal controls. In this instance the light stimuli were presented by a strip light on the front panel of the apparatus which came on for 6 sec. or for 2 sec. directly in front of the subject. The touch stimulus in both these transfer series was administered in the same way as in the learning task. However, while the learning stimulus was a hard-ended rod, the transfer stimulus was a soft brush. At the end of the session, each child was asked: 'How did you know when to knock and when to rub?' All answers were noted. For the deaf children, the question was 'signed' in addition to being asked verbally. Any adequate signing, demonstration or attempt at verbalization was accepted.

At the end of the transfer session retention of original learning was checked on 18 trials of the original stimulus. Anyone falling short of the learning criterion was rejected.

As a check on stimulus equivalence four additional groups, of 10 normal children each, were presented with 18 touch, air, light or sound stimuli and were asked to discriminate between them in the same way as in the learning task in the main experiment. Errors for discrimination at each modality were scored.

RESULTS

The results are presented in two parts: (1) learning error scores on the initial tactile task using the hard touch stimulus, (2) transfer error scores.

Learning error scores

All subjects learned to a criterion of 51 out of 54 trials of consecutive responses provided these did not exceed 90 trials. The number of errors made by each group in achieving criterion are given in Table 2.

Table 2. *Error scores to learning criterion by groups*

	Blind	Normal	Subnormal	Deaf	Normal	Subnormal
Mean	1·8	8·6	3·7	5·8	7·3	4·2
S.D.	1·33	7·42	2·53	2·95	4·29	3·37

Analysis of variance of these error scores showed a significant group difference ($F = 3{\cdot}31$; d.f. 5, 54; $P < 0{\cdot}025$). The blind had fewer learning errors than other groups. The two subnormal control groups also had fewer errors. The deaf had more errors than the subnormals and the normal children had most learning errors. Correlations were run for all subjects between error to criterion on the learning task and transfer error scores. No correlation reached significance.

The transfer error scores were analysed in several ways. Transfer can be defined as a saving score based on the difference between the first 18 learning trials on the learning task and the 18 trials on any transfer task. Using this technique, difference scores were tested on *A* tests (Sandler, 1955) for blind children and the appropriate control groups and for deaf children and the two appropriate controls. For each group, three tests were carried out, learning errors *v.* touch transfer errors, learning errors against air-puff transfer errors, and learning errors against light transfer errors, or against sound transfer errors. Only two of these tests were significant and these were for the deaf children who had fewer errors in air-puff and touch transfer than the initial touch learning trials. On the basis of these error scores comparisons, it must be concluded that transfer did not occur crossmodally in any group and only in one of the six groups did it occur within a dimension (air-puff) and within scale (touch).

Another definition of transfer might be the difference between a criterion learning score, in this case zero errors in the final 18 trials of learning, and transfer error scores. This is a more stringent test and involves testing the significance of the difference of transfer error scores from zero. Criterion learning error scores (the last 18 trials) were compared with the transfer error scores in each condition and each group by similar *A* tests. No results showed any transfer scores to be equal to criterion learning error scores. Mean transfer error scores are presented in Table 3.

Table 3. *Transfer error score by groups and conditions*

(Means and SDs)

Conditions	Blind	Normal	Subnormal
Touch	2·7 ± 3·4	2·7 ± 3·6	2·3 ± 3·3
Air	3·9 ± 3·6	4·5 ± 4·1	3·5 ± 3·8
Sound	3·6 ± 3·7	4·2 ± 3·9	3·0 ± 3·6
	Deaf	Normal	Subnormal
Touch	0·6 ± 1·2	4·1 ± 3·2	2·1 ± 3·0
Air	1·2 ± 1·1	3·0 ± 3·1	3·1 ± 2·7
Light	3·5 ± 5·1	3·7 ± 3·7	4·4 ± 3·3

The transfer error scores were analysed against themselves in two mixed analyses of variance designs for the blind and controls, and for the deaf and controls. In both cases a significant conditions term resulted ($F = 6·31$, d.f. = 2, 54, $P < 0·01$, and $F = 4·60$, d.f. = 2, 54, $P = < 0·025$, respectively). A final combined analysis gave the same result ($F = 6·43$; d.f. = 2, 108; $P < 0·001$). Appropriate t tests showed that transfer was greatest to touch and least to light and sound. Air-puff error scores were intermediate.

Finally the discrimination learning scores of the additional normal groups who learned the transfer tasks without an initial learning task were analysed. No differences were found between the number of errors in touch, air-puffs and light. These tasks were therefore equivalent in difficulty. Sound, however, proved to be significantly easier to learn. This difference is not reflected in greater transfer.

The number of children from each group who gave correct answers to the question of how they knew which response to give are shown in Table 4.

Table 4. *Percentage correct verbalization*

Groups	Percentage
Blind	90
Deaf	40
Subnormal controls (for blind group)	90
Subnormal controls (for deaf group)	80
Normal controls (for blind group)	60
Normal controls (for deaf group)	70

These frequencies within any group were unrelated to the amount of transfer, but did indicate a non-significant relationship with the efficiency of the initial discrimination learning.

DISCUSSION

The initial temporal touch discrimination task proved difficult for the subjects. This can be seen by the rejections necessary to find 10 children in each group who could do the task. However, as the learning and the transfer tasks were of similar difficulty when each was given as a learning task, the lack of transfer cannot be attributed to the greater difficulty of discrimination other than touch stimuli. In fact, one of the transfer tasks, the discrimination of sound durations, was easier to learn than the tactile discrimination.

The basic result of this study is that with the procedure used, few children show evidence of crossmodal transfer as defined in this experiment. Only the deaf show a tendency to successful transferring to touch and to air-puffs. One conclusion which can be drawn from this is that adequate verbalization is not a sufficient condition for crossmodality transfer to occur. Adequate verbalization occurred in most of the groups whose members in the experiment failed to transfer, and the inadequacy of speech in the deaf was no special handicap in comparison with other groups. A second conclusion of equal importance is that general intellectual impairment, presumably accompanied by general diffuse brain damage, is not an additional handicap to crossmodal transfer where mental age is matched with that of other groups.

Concerning adequate verbalization combined with the lack of transfer found in most groups in this experiment, Bryant (1967) has noted that at a certain level of ability speech seems to be specific to the task learned, and does not generalize. Apparently in our experiment words acted as specifically as do other motor responses, i.e. they adequately described the principle of solution of one task, but were not seen as generalizable to similar tasks in other sense modalities.

Curiously perhaps, blindness seems to be more disabling than deafness for intermodal transfer, a fact which may point to the relevance of sensory integration as discussed by Held & Hein (1963) rather than to the importance of language. It also supports Piaget's view noted in the introduction. In the case of the blind, the initial discrimination learning was faster than in the other groups, so that their failure to transfer cannot be attributed to any lower level of cognitive ability. Nine of the 10 blind subjects also verbalized the principle involved in their learning task.

It is interesting that air-puffs which were assumed to be on a similar dimension to the tactile stimuli were treated as equivalent to light and sound stimuli in the children's responses. This makes the arbitrary division into unimodal and cross-modal transfer somewhat suspect. In future studies, perhaps, such boundaries should be determined empirically rather than *a priori*.

It may be remarked that when learning error scores and transfer error scores are examined, they tend to be less than would occur by chance. In this study, therefore, as in many others, transfer is measured by a saving score, but learning is always relatively advanced at an early stage in both the initial and the transfer task.

The results show that in the situation examined and using a saving score as a measure of transfer, intermodal transfer does not occur, despite the frequent verbalization of the principle involved.

We are grateful for the help given us by the superintendent, staff and patients of St Lawrence's Hospital, Caterham, and by the head teachers, staff and pupils of Ackmar Road School, Dog Kennel Hill School, Dorton House, Elizabeth Burgwin, Elm Court Road School, Linden Lodge, Meadway, St Christopher's and Waldron Road School.

REFERENCES

Birch, H. G. & Belmont, L. (1965). Auditory-visual integration in brain-damaged and normal children. *Dev. Med. Child Neurol.* **7**, 135–144.

Birch, H. G. & Lefford, A. (1963). Intersensory development in children. *Monogr. Soc. Res. Child Dev.* **28**, no. 5, 1–48.

Blank, M., Altman, L. D. & Bridger, W. H. (1968). Crossmodal transfer of form discrimination in pre-school children. *Psychon. Sci.* **10**, 51–52.

Blank, M. & Bridger, W. H. (1964). Crossmodal transfer in nursery school children. *J. comp. physiol. Psychol.* **58**, 277–282.

Blank, M. & Bridger, W. H. (1966). Conceptual crossmodal transfer in deaf and hearing children. *Child Dev.* **37**, 29–38.

Bryant, P. E. (1967). Verbal labelling and learning strategies in normal and severely subnormal children. *Q. J. exp. Psychol.* **19**, 155–161.

Ettlinger, G. (1967). Analysis of crossmodal effects and their relationship to language. In C. H. Millikan & F. L. Darlep (eds.), *Brain Mechanisms Underlying Speech and Language.* New York: Grune & Stratton.

Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain* **88**, 237–294, 585–644.

Held, R. & Hein, A. (1963). Movement produced stimulation in the development of visually guided behaviour. *J. comp. physiol. Psychol.* **56**, 872–876.

Hermelin, B. & O'Connor, N. (1964). Crossmodal transfer in normal, subnormal and autistic children. *Neuropsychologia* **2**, 229–235.

Luria, A. R. (1961). *The Role of Speech in the Regulation of Normal and Abnormal Behaviour.* Oxford: Pergamon Press.

Merlin, V. S. (1957). The dynamics of 'transfer' of conditioned reflex connections from one signal system to another. *Vopr. Psikhol.* **2**, 53–67.

O'Connor, N. & Hermelin, B. (1961). Like and crossmodality recognition in subnormal children. *Q. J. exp. Psychol.* **13**, 48–52.

Piaget, J. (1965). Langage et pensée. *Rev. Practicien* **15**, 2253–2254.

RUDEL, R. & TEUBER, H.-L. (1968). Pattern recognition within and across sensory modalities in normal and brain-injured children. (Annual Report, Massachusetts Institute of Technology, Department of Psychology.)

SANDLER, J. (1955). A test of the significance of the difference between the means of correlated measures, based on a simplification of Student's *t*. *Br. J. Psychol.* **46**, 225–226.

SMITH, J. & TUNICK, J. (1969). Transfer of discrimination by retarded children. *J. exp. Child Psychol.* **7**, 274–281.

STEVENS, S. S. & GALANTER, E. H. (1957). Ratio scales and category scales for a dozen perceptual continua. *J. exp. Psychol.* **54**, 377–411.

Manuscript received 2 July 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 355–359
With 1 text-figure
Printed in Great Britain

Symptom Assessment in Epidemiology: A Comparison between Two Methods

BY M. M. WOOD

Medical Research Council, Llandough Hospital, Penarth, Glam.

A self-administered form of subjective symptom assessment is described, based on a method evolved by Ingham. The self-administered scales provided data in a form which could be collected easily and quickly without the necessity of a skilled interviewer. Although the information required decoding, this could be carried out by clerical staff under supervision. When compared with Ingham's individually administered form, there were indications that the self-administered scales produced slightly more inconsistent responses and for this reason were sometimes more difficult to score. Further investigations using the self-administered scales after slight modification have achieved notable success, especially where there has been some supervision.

Epidemiological surveys with individual interviewing are expensive, particularly when home visits are involved. Whether it is possible to obtain the same information by cheaper means must depend upon a number of factors that vary from survey to survey, in particular the sort of information that is required. In the course of a collaborative project, which for administrative reasons had to be completed quickly (Wood & Elwood, 1966), an attempt was made to compare individual interviewing with a self-administered questionnaire, completed without supervision. The survey was a study of anaemia and the information sought by questionnaire was a numerical assessment of some relevant symptoms.

One technique for the assessment of symptoms in epidemiology was described by Ingham (1965, 1968) and requires individual administration. Similar scales were devised in both individual and self-administered versions and answers were sought to the following questions: (i) How much time was saved by using self-administered scales rather than individually administered scales? (ii) How reliable was the information obtained in this way compared with the individually administered form? (iii) Would the same results have been obtained if the individual form had been used?

METHOD

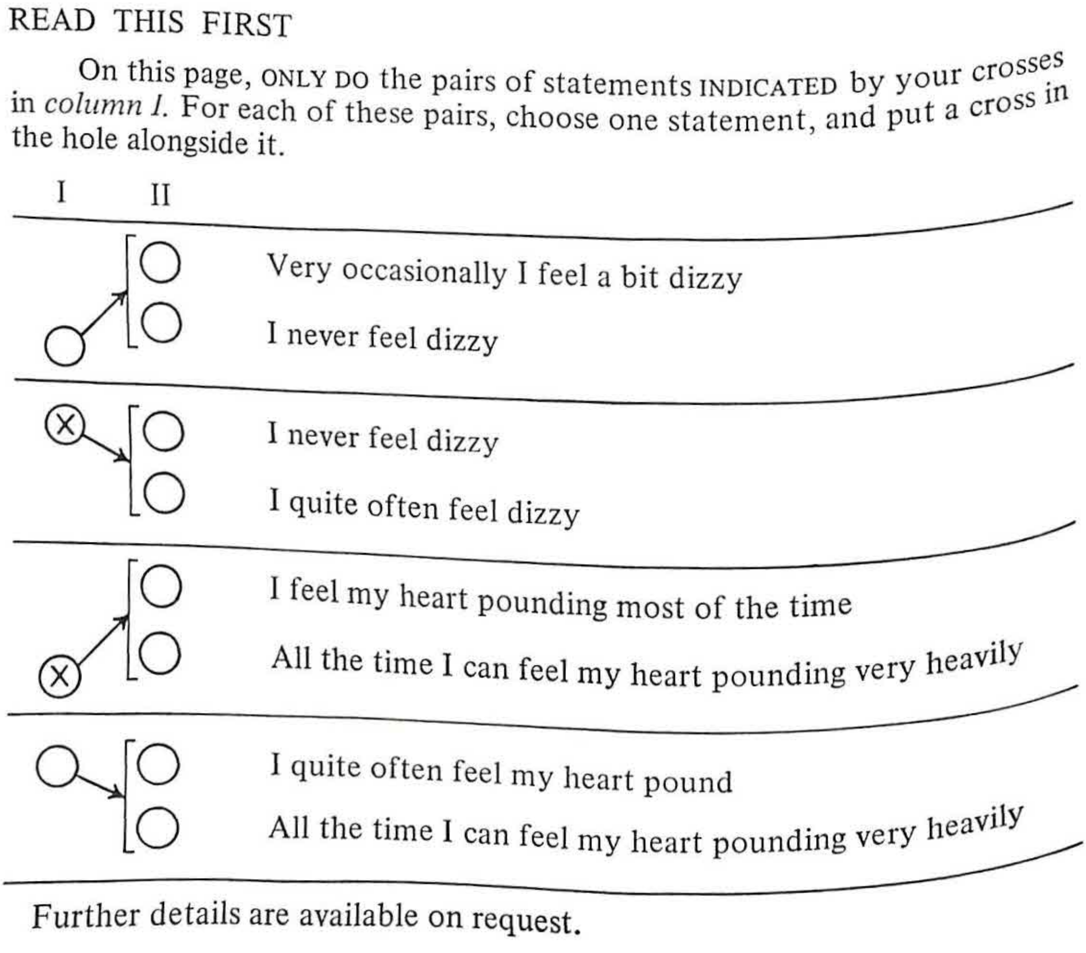
Sample. The whole population between the ages of 15 and 74 years, living within a quarter of a mile of one specified local health authority clinic in South Wales, was asked to cooperate in the main survey. The number agreeing was 295 and these comprised the sample for the present investigation. The purpose of the main survey was to study anaemia, and the symptoms to be assessed were chosen accordingly. They were: irritability, palpitations, dizziness, breathlessness, fatigue, and headache. A subsample comprising all females aged between 35 and 55 years ($n = 37$) was selected for individual interviewing at home.

Individual symptom scales. Each scale was defined by a set of seven statements representing different degrees of severity. The statements used for the 'fatigue' scale, for example, were:

(1) I never get tired. (2) I only feel tired after a very hard day. (3) I generally feel a little tired by the end of the day. (4) I always feel very tired by the end of the day. (5) I feel very tired most of the time. (6) I always feel very tired all day long. (7) I am always so worn out and exhausted, I feel I never want to do anything.

The statements were presented to the subject in pairs and for each pair he was asked to indicate which statement was nearer to the truth as far as he was concerned. Eleven such pairs, each printed on a separate card, were available for each scale. Some economy of presentation was achieved by presenting them in three stages, the second and third stages involving a selection by the investigator on the basis of the previous responses. Two of the scales, 'fatigue' and 'headache', had been used previously but the stages of presentation were different (Ingham, 1965).

Self-administered symptom scales. It would be simple to print the statements in pairs on several pages of a booklet and to ask subjects to work through them in exactly the same way as they would if they were presented on separate cards. This would entail 66 pairs of statements for each subject, however, and a method was sought for reducing this number by 'programming' the procedure as in the individual scales. This was done by an arrangement in which the mark made by the subject, in response to each pair, also appeared on the following page and indicated which pair he was to consider there. The subject was asked to mark his choice by placing a cross in the hole or circle alongside each statement. The holes allowed the cross to appear on the next page, indicating the pairs of statements to be done next. An example of what a subject might see on turning a page of the booklet is shown in Fig. 1.



READ THIS FIRST

On this page, ONLY DO the pairs of statements INDICATED by your crosses in *column I.* For each of these pairs, choose one statement, and put a cross in the hole alongside it.

I	II	
○	○	Very occasionally I feel a bit dizzy
	○	I never feel dizzy
☒	○	I never feel dizzy
	○	I quite often feel dizzy
☒	○	I feel my heart pounding most of the time
	○	All the time I can feel my heart pounding very heavily
○	○	I quite often feel my heart pound
	○	All the time I can feel my heart pounding very heavily

Further details are available on request.

Fig. 1. Sample of page 2 from self-administered symptom scales.

Items from all the symptom scales were mixed throughout the booklet, and the order was randomized as far as possible, bearing in mind the limitation due to programming. The order of statements within pairs was also randomized. A partially worked example was given on the first page of the booklet to familiarize subjects with the procedure.

Procedure. The booklets were given to people as they left the clinic after haematological and dental examinations. They were asked to complete and return the booklets within 14 days, in the stamped addressed envelope provided. Home visits were made to interview the subsample and to collect questionnaires not returned, or completed incorrectly.

RESULTS

The average time taken by a psychologist to interview a subject and administer the individual symptom scales was 30 min. The decoding of the individual symptom scales took place during the interview. The self-administered scales were handed to subjects by a clerical assistant with a brief word of explanation, taking about 1 min.; the average time taken to decode the responses afterwards was 9 min., including checking. Decoding of the self-administered scales could have been carried out by clerical staff, but the interviewing required a skilled research worker.

Although 71 per cent of those people who attended the clinic returned their questionnaires spontaneously, a home visit was necessary to the remaining 29 per cent to collect the questionnaires. Of the people from whom questionnaires were eventually obtained (95 per cent) home visits were made to 19 per cent in order to correct questionnaires. A rough estimate of the average time required for collecting a questionnaire from a subject who failed to return it is 10 min. When it was necessary to correct it as well it took approximately 15 min. per subject. The correcting had to be done by a psychologist because it involved the same skills as those required for individual testing. Averaged over the complete sample, therefore, home visits entailed an extra 3 min. per subject of work that could have been done by non-graduate assistants, plus 3 min. of psychologist's work. The total of 16 min. per subject, mostly involving relatively unskilled work, compared with at least 30 min. of skilled work per subject for the individual scales represents a substantial saving for the self-administered form. It took a subject about 20 min. to complete the self-administered scales when under observation. It is possible, therefore, that there was an overall saving of time for subjects as well as research workers, though there is no doubt that many prefer the friendliness of the individual contact.

Of the subsample chosen to be given the individual symptom scales, 18 were interviewed before they attended the clinic (group I), while the other 19 were interviewed after they had returned their booklets (group II).

An index of unreliability of the scales is the rate of inconsistent responses. This is defined as the number of inconsistent responses divided by the number of inconsistent responses possible without changing the scale position. This definition tends to inflate the inconsistency rate slightly. The individual scales produced a lower rate of inconsistent responses (3·5 per cent) than the self-administered scales (8·0 per cent), and there seemed to be a small beneficial effect on the self-administered form when the individual scales were given first.

The correlations between individual and self-administered scales are shown in Table 1. Only one correlation is not significant, for 'breathlessness' in group II.

A contributory factor may have been that the statements for 'breathlessness' were not prepared with the same care as the statements for 'fatigue' and 'headache'. In addition, more inconsistent responses were produced on 'breathlessness' (15) than on any of the other five symptoms (mean 4·6 for group II).

Table 1. *Correlation coefficients (r) between scores on individual and self-administered scales*

Symptom	Group I ($n = 18$)	Group II ($n = 19$)
Irritability	0·69	0·86
Palpitations	0·74	0·69
Dizziness	0·63	0·58
Breathlessness	0·57	0·41
Fatigue	0·84	0·83
Headache	0·83	0·83

($r > 0{\cdot}44$; $P < 0{\cdot}05$)

DISCUSSION

It is clear that the individual symptom scales are more reliable, in terms of inconsistent responses, than the self-administered scales. There were also fewer problems in ascribing scores to response patterns. In the individual interview situation, difficulties, ambiguous response patterns, and problems of cooperation can be dealt with easily. Also interference by other members of the family or friends, which is sometimes a problem with self-administered scales and questionnaires, can be prevented.

However, the self-administered scales saved time, and did enable the data to be collected very quickly and processed at a more convenient time. In this survey the time and staff available for interviewing were limited. The conversion of the responses into patterns and scores could have been carried out by clerical staff under supervision.

Some of the disadvantages of the self-administered scales have been overcome in subsequent surveys by simplifying the format, at the price of some loss of discriminatory power, and also by ensuring that they were completed under supervision (Elwood *et al.*, 1967). Both of these developments have ensured an improvement in the consistency of responses.

The author would like to thank Dr J. G. Ingham and Dr P. C. Elwood for their invaluable advice and assistance.

REFERENCES

ELWOOD, P. C., WATERS, W. E., GREENE, W. J. & WOOD, M. M. (1967). Evaluation of a screening survey for anaemia in adult non-pregnant women. *Br. med. J.* iv, 714–717.

INGHAM, J. G. (1965). A method for observing symptoms and attitudes. *Br. J. soc. clin. Psychol.* 4, 131–140.

INGHAM, J. G. (1968). The assessment of symptoms for clinical and epidemiological purposes. *Proc. XVI Int. Congr. Appl. Psychol.* 702–706.

WOOD, M. M. & ELWOOD, P. C. (1966). Symptoms of iron deficiency anaemia: a community survey. *Br. J. prev. soc. Med.* **20**, 117–121.

Manuscript received 15 *July* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 360–361
Printed in Great Britain

The Performance of Elderly Subjects on Gooddy's Arrow Drawing Test

By GEORGE IRVING

Department of Psychological Research,
Crichton Royal, Dumfries

Gooddy & Reinhold (1964) described a simple test which they claimed may be of value in detecting early or slight disease of the brain, especially of the minor or non-dominant hemisphere. The test consisted of asking patients to draw arrows in a number of specified directions. The present report describes results obtained from the administration of the test to a group of elderly psychiatric patients, a group of 'normal' elderly subjects, and a group of psychiatric patients under the age of 65 years.

The elderly psychiatric group consisted of 54 patients; 27 were diagnosed as either senile or arteriosclerotic dementias (mean age 76·0 years, S.D. 7·0) and the majority of the remainder were classified as depressions (mean age 72·2 years, S.D. 6·6). The normal group consisted of 40 subjects who had been selected at random from the lists of local general practitioners and were investigated as part of a wider social-psychiatric investigation. Seven were, in fact, considered to be showing evidence of psychiatric abnormality (three organic, four functional). The mean age of this group was 73·9 years (S.D. 5·9). The psychiatric group under 65 years consisted of 54 approximately consecutive hospital admissions (mean age 41·0 years, S.D. 12·1); none was considered to be suffering from any form of brain damage. Patients in the elderly psychiatric group and in the normal elderly group were, in addition to the arrow test, given a short battery of psychometric tests of intellectual impairment.

The method of administration of the arrow test was altered slightly from that described by Gooddy. An arrow (→) was drawn on a sheet of paper in full view of the subject with the instructions, 'This is an arrow pointing to this side of the paper'. The model was then removed and the subject asked to draw arrows in each of six specified directions. No comment was made on the quality of any drawing.

RESULTS

1. In the elderly organic group 20 patients made errors on the test. Of these, nine made miscellaneous errors which were not readily classifiable. Eleven patients made a specific type of error on one or more of their drawings. This specific error was that of reversing the head of the arrow, i.e.—<. In the elderly functional group, 10 patients made this reversal error; none made errors of any other kind. In the normal elderly group 11 subjects made errors; of these 10 reversed the arrow-head and one made errors of a different kind; only one of the seven subjects showing evidence of psychiatric disorder performed abnormally on the test. (This subject was diagnosed as a mild dementia.) In the under-65 group no patients made errors of any kind.

2. No significant relationship was found between the occurrence of errors and tests of intellectual impairment in either the functional elderly or normal elderly groups. Neither was there any relationship between errors and position above/below the median age in either group.

3. Since this study was carried out a further 56 elderly psychiatric patients (28 organic and 28 functional) have been given the test and approximately the same pattern of abnormal performance has been observed in both diagnostic groups.

CONCLUSION

The main feature of this study is the relatively high incidence of arrow-head reversals in the normal elderly and functional elderly groups. However, on the basis of the evidence available it is not possible to evaluate the significance of this observation.

REFERENCE

GOODDY, W. & REINHOLD, M. (1964). The sense of direction and the arrow form. In L. Halpern (ed.), *Problems of Dynamic Neurology*. Jerusalem: Hebrew University, Hadench Medical School.

Manuscript received 7 July 1970

Revised manuscript received 3 September 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 362–366
Printed in Great Britain

A Comparative Study of Criminals and Matched Controls on Three Dimensions of Personality

By SYBIL B. G. EYSENCK AND HANS J. EYSENCK

University of London

Six hundred and six male trainee railmen and 518 criminals were administered personality inventories purporting to measure the personality dimensions of Psychoticism (P), Extraversion (E) and Neuroticism (N). Also administered was a Lie scale (L). Scale means and reliabilities are reported for the experimental and control groups, and proportions of 'Yes' answers compared for individual items. The prisoners are significantly higher on P and N, but lower on E, than the controls; this may be due to the absence of impulsiveness items in the E scale used, and the preponderance of sociability items. There are no differences on the L scale, suggesting that prisoners did not attempt to fake scores. The groups were matched for age and social class, as well as sex, all being male. Coloured members originally forming part of the control group were found to differ significantly from the white members, and were removed prior to the comparisons being run. The results found are on the whole similar to previously reported data using different inventories and different comparison groups, and a different sample of prisoners.

In his book *Crime and Personality*, Eysenck (1964) has developed a series of theories relating criminal behaviour to various dimensions of personality. In particular, he suggested that criminals would score high on measures of N (neuroticism), P (psychoticism) and E (extraversion). The literature reviewed in that book, and papers apprearing since (Sanocki, 1969; Eysenck & Eysenck, 1970, 1971), have on the whole supported these hypotheses, although deductions from the theory have not always been supported by the data at a high level of significance. In addition, it has become apparent that the prediction relating to E was correct only in part; Eysenck & Eysenck (1971) have shown that items measuring the sociability component of E were not positively related to criminality (and might even be negatively related), while items measuring impulsiveness were positively correlated with criminality. This finding could be interpreted to support the original hypothesis linking E with criminality through the mechanism of *conditioning* because eye-blink conditioning, too, was found to be correlated with E (negatively) only by way of impulsivity, not by way of sociability (unpublished). This important division of E into two main constituent parts has been discussed in detail by Eysenck & Eysenck (1963).

The experiment reported here is in some ways a replication of the one discussed by Eysenck & Eysenck (1970, 1971), in which an experimental group of criminals was given a personality inventory measuring P, E and N, and their scale scores and item endorsements compared with those of three control groups (random group of parents, students and apprentices). None of the control groups was entirely satisfactory, and it was hoped that their imperfections might cancel out when detailed comparisons were made. In the present experiment a more carefully matched control group is used in making comparisons with a sample of prisoners chosen along the same lines as in our previous experiment. In addition, a Lie scale (L scale) is incorporated in the personality questionnaire used, in order to test the hypothesis that the prisoners may be faking good (or bad); the rationale of such scales is discussed in detail by Michaelis & Eysenck (1971). The items used for the P, E and N scales in the present inquiry are largely different from those used in our previous one; they are in fact identical with those used in our factorial study aimed at demonstrating the identity of P as a separate dimension of personality (Eysenck & Eysenck, 1968). In this scale (called the PEN scale for short) the E factor is represented almost entirely by sociability items, and not at all by impulsiveness items; hence our prediction for the present study would have to be that criminals would not score more extraverted than controls, but would either have similar scores or have even lower ones.

Our criminal population was made up of 606 male criminals with a mean age of 26·53 (S.D. 10·00); these were a random population from several of the main British prisons. Our control population consisted of 518 white trainee railmen employed by London Transport, all male and with a mean age of 25·85 (S.D. 10·69). The two samples are thus reasonably well matched on sex, age and social class; they differ in that all coloured members had been removed from a somewhat larger total sample of controls. In all, 140 of these coloured trainee railmen were removed because they were found to have significantly different scores on some of the scales; details regarding these differences will be given presently. Coloured persons could not be removed from the prison sample because no record was available to identify them by colour; in any case, the proportion of coloured to white prisoners would almost certainly be smaller. The age of the coloured controls was 30·83 (S.D. 9·42).*

Table 1. *Means and S.D.s of* 606 *prisoners, and* 518 *white and* 140 *coloured London Transport employees on P, E, N and L scales*

			London Transport employees			
	Prisoners		White		Coloured	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
P	4·81	3·95	3·14	2·85	3·74	2·50
E	12·43	3·95	13·58	3·59	13·58	2·83
N	9·56	4·46	7·53	4·40	6·38	3·83
L	5·48	3·32	5·44	3·26	8·96	3·63

Table 1 gives the means and S.D.s of these various samples for P, E, N and L. It will be seen that prisoners are significantly higher than controls on P and N, and significantly lower than controls on E; there are no differences on L, suggesting (if we are willing to assume that L is at least in part a measure of dissimulation) that deliberate falsification did not play an important part in the production of these differences. Coloured people are much higher on the L scale than white controls; they are also higher on P ($P < 0{\cdot}05$) and lower on N ($P < 0{\cdot}01$). There are no differences on E.

---

* Unselected inventories were received from London Transport, and the allocation of the groups was undertaken by us.

The need to match samples for social class is brought out clearly if we compare the data given in Table 1 with the norms for a more representative population sample of 1012 male subjects. These score lower on P (2·50; S.D. 2·71) and on E (12·75; S.D. 4·12) than our working-class white controls; on N they score only fractionally lower (7·33; S.D. 4·37.) These results are in reasonably good agreement with the social class comparisons made in our previous paper (Eysenck & Eysenck, 1969). On L we have available results from two samples: a random sample of 152, and a somewhat more middle-class one of 329. L scores were 4·56 (S.D. 2·95) and 3·56 (S.D. 2·48) respectively, suggesting that L scores are higher in working-class groups. This does not necessarily mean that working-class subjects lie more; alternative and possibly more likely reasons for this difference are given in Michaelis & Eysenck (1971).

Analyses of item comparisons are reported in Tables 2–4 for P, E and N, respectively. Only significant differences are reported, and for this purpose all differences above 5 per cent have been used. A number of smaller differences might have

Table 2. *P scale items showing significant differences between criminals and controls*

	Percentage Yes answers	
	Criminals	Controls
Have you had an awful lot of bad luck?	67	31
Have you have more trouble than most?	54	20
Would you have been more successful if people had not put difficulties in your way?	60	38
Is there someone who is responsbile for most of your troubles?	36	21
Do you worry a lot about catching diseases?	28	15
Are there several people who keep trying to avoid you?	23	12
Do you have enemies who wish to harm you?	23	14
Do your friendships break up easily without it being your fault?	24	17
Do you let your dreams warn or guide you?	20	13
Would you take drugs which may have strange or dangerous effects?	18	12
Was your father a good man?	79	85

Table 3. *E scale items showing significant differences between criminals and controls*

	Percentage Yes answers	
	Criminals	Controls
Do you find it hard to show your feelings?	56	40
Would you call yourself happy-go-lucky?	51	64
Are you more distant and reserved than most people?	40	30
Can you usually let yourself go and enjoy yourself a lot at a gay party?	77	85
Would you call yourself talkative?	35	42
Do you nearly always have a 'ready answer' when people talk to you?	38	45
Do you like plenty of bustle and excitement around you?	61	68
Do you like mixing with people?	81	88
Do you like people around you?	81	88
Do you normally prefer to be alone?	31	23
Do you make friends easily with members of your own sex?	80	87
Would you enjoy hunting, fishing and shooting?	74	69

qualified because the proportion of endorsements was near the 0 per cent or 100 per cent level, but such differences are too small anyway to be of any psychological interest. In order to calculate the proportions and differences, the sample of prisoners was reduced to the same size as that of the controls ($n = 518$) by a random process.

Table 4. *N scale items showing significant differences between criminals and controls*

	Percentage Yes answers	
	Criminals	Controls
Are you often troubled about feelings of guilt?	57	33
Have you ever been afraid of losing your mind?	48	26
Have you ever wished you were dead?	47	25
Do you often feel fed up?	70	54
Do you feel self-pity now and again?	66	50
Do you sometimes feel uneasy indoors?	58	42
Do things sometimes seem as if they are not real?	68	51
Do you get depressed in the mornings?	40	28
Do you sometimes feel you don't care what happens to you?	52	40
Do you often feel very weak all over?	37	26
Do you find it hard to keep your mind on what you are doing?	34	25
Do you worry a lot about your looks?	41	32
Do you ever feel 'just miserable' for no good reason?	50	43

These item analyses bear out our expectations, and also of course the comparisons of sample means; prisoners are more likely to give answers indicative of psychoticism, of neuroticism and, to a lesser extent, of lack of sociability. The particular items quoted are similar to those which in our previous study differentiated prisoners and controls, although of course the wording of many of these items is not identical. One possible criticism should perhaps be anticipated here. It might be argued that some of the items showing differences between prisoners and controls might be influenced by the circumstances and status of the prisoners rather than by their personalities. Thus the answer to the question: 'Have you had an awful lot of bad luck?' might be thought to reflect the prisoner's reaction to being caught and sent to prison rather than his habitual personal feeling. Such an objection must be granted some force, and almost certainly the circumstances of being in prison have increased the differences observed on some of the questions between the two groups. But it is not likely that this is the whole story; we have shown that for both prisoners and controls similar factors emerge from factor analyses of such questionnaires (Eysenck & Eysenck, 1970), and many of the questions related to each factor cannot conceivably be influenced in this manner (e.g. 'Do you let your dreams warn or guide you?' 'Would you enjoy hunting, fishing and shooting?' 'Do you worry a lot about your looks?'). We have suggested elsewhere (Eysenck & Eysenck, 1970) the use of a 'criminality scale' made up of the most diagnostic items in order to make parole and other predictions; this would appear to be the best way of testing the assumptions underlying objections of this kind. Another possibility would of course be the use of follow-up studies of adolescents to whom questionnaires would be administered before their criminal careers started; it is very much to be hoped that such studies will in due course be undertaken. *Ex post facto* investigations, such as the one here described, always suffer from weaknesses which cannot in the nature of the case be overcome. Such prospective studies as have been published have rather strongly supported the hypotheses here advanced (Eysenck, 1964).

We are indebted to the Maudsley and Bethlem Royal Research Fund for the support of this investigation. Dr P. A. B. Raffle, Chief Medical Officer of London Transport, kindly gave permission for the testing of the trainee railmen, and we would especially like to thank Mr R. C. Potter, Chief Instructor, Railway Training Centre, who arranged for the questionnaire to be administered. We are also indebted to the Home Office for permission to test prisoners; most of the testing was carried out by prison psychologists to whom we wish to express our gratitude for their generous help. Opinions expressed in this article are of course our own, and do not necessarily represent the view of the Home Office or London Transport.

REFERENCES

EYSENCK, H. J. (1964). *Crime and Personality*, 2nd ed. London: Paladin Books, 1970.

EYSENCK, S. B. G. & EYSENCK, H. J. (1963). On the dual nature of extraversion. *Br. J. soc. clin. Psychol.* **2**, 46–55.

EYSENCK, S. B. G. & EYSENCK, H. J. (1968). The measurement of psychoticism: a study of factor stability and reliability. *Br. J. soc. clin. Psychol.* **7**, 286–294.

EYSENCK, S. B. G. & EYSENCK, H. J. (1969). Scores on three personality variables as a function of age, sex, and social class. *Br. J. soc. clin. Psychol.* **8**, 69–76.

EYSENCK, S. B. G. & EYSENCK, H. J. (1970). Crime and personality: an empirical study of the three-factor theory. *Br. J. Criminol.* **10**, 225–239.

EYSENCK, S. B. G. & EYSENCK, H. J. (1971). Crime and personality: item analysis of questionnaire responses. *Br. J. Criminol.* **11**, 49–62.

MICHAELIS, W. & EYSENCK, H. J. (1971). The determination of personality inventory factor patterns and intercorrelations by changes in real-life motivations. *J. genet. Psychol.* **118**, 223–234.

SANOCKI, W. (1969). The use of Eysenck's inventory for testing young prisoners. *Przeglad Penitencjarny, Warszawa* **7**, 53–68.

Manuscript received 16 July 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 367–374
With 3 text-figures
Printed in Great Britain

An Ordinate Comparison Method of Calculating Brain Damage Probability

By BRIAN R. GLAISTER

Netherne Hospital, Coulsdon, Surrey

In clinical practice it is frequently more useful to calculate a numerical value of brain damage probability prior to, or as an alternative to, classification. A method of calculating brain damage probability is described which is a comparison of the proportions in the brain-damaged and non-brain-damaged groups which obtain the given score. The method compounds national norms and a local base rate. It could be described as an ordinate comparison method. The validity of a calculated brain damage probability depends, *inter alia*, on the score, the base rate, the method of calculating the probability, the reliability of the base rate and the reliability of the national norms. An efficiency function which is a function of the score, the base rate and the method of calculating the probability, is calculated: methods of estimating the reliability of the base rate and of the national norms are given. The efficiency of brain damage probability calculated by this ordinate comparison method compares favourably with other methods.

Published norms for brain damage tests are usually either based on equal-sized control and experimental groups, e.g. the Revised Visual Retention Test (Benton, 1955), or are mathematically equalized by conversion to percentiles prior to use in test score evaluation, e.g. the Memory-for-Designs Test (Graham & Kendall, 1960). Subsequent use of such data in clinical work assumes a base rate of 50 per cent brain-damaged, which is an unreasonable assumption for most clinical populations. The example presented here makes use of the brain damage base rate of 29 per cent found at Netherne for 1967.

The example also makes use of the Modified Word Learning Test (Walton & Black, 1957). The norms have been obtained by combining two sets of published data: Walton & Black (1957) and Walton *et al.* (1959, excluding subnormals and ESN). There are 124 subjects in the brain-damaged group and 393 in the non-brain-damaged group. The scores are calculated by the revised method (Walton *et al.*, 1959).

Published norms for brain damage tests usually divide the scores into a few categories—sometimes into just two, as is the case for the Modified Word Learning Test. In clinical practice it is more often useful and reasonable to use the normative data to calculate numerical values of brain damage probability. The method given here is one of the methods of doing this.

Methods of calculating score abnormality usually consider not only the score obtained by the examinee but also all scores indicating an even higher likelihood of brain damage. The method given here, however, considers only the score obtained by the examinee; all other scores are disregarded. For this reason it could be described as an ordinate comparison method.

The method presented can be derived from Bayes's theorem (Bayes, 1958). But, as Kendrick (1965) has pointed out, Bayesian statistics are for this purpose 'common sense' statistics and nothing would be added to the clarity of presentation by using Bayes's notation.

THE BASE RATE

A follow up of those referred for testing at Netherne during 1967 as possible brain-damage cases showed that 29 per cent had been diagnosed as such 18 months later. This was conventional psychiatric diagnosis in which only a few cases could be assigned with absolute certainty. The base rate is thus only as reliable as the psychiatric diagnosis, there being some evidence for reliability being higher for the diagnosis of organic conditions than for other broad categories (Kreitman, 1961; Beck, 1962). However, even if unreliable, this base rate can be accepted with greater confidence than the 50 per cent otherwise assumed. The data for the Netherne base rate are given in Table 1.

Table 1. *The composition of the group referred for testing at Netherne during 1967 as possible brain-damage cases*

	n	%
Brain-damaged	12	29
Not brain-damaged	29	71
Not known	5	—
Total	46	100

This base rate is probably similar to that found in most general mental hospitals but greater precision will, of course, result from the ascertainment of the base rate for the situation in which the test is to be used.

THE ORDINATE COMPARISON CALCULATIONS

Consider the published norms. Let p_1 be the percentage of the brain-damaged group which obtains score x and let p_2 be the percentage of the non-brain-damaged group which obtains score x. A comparison of these two percentages would give a measure of brain damage probability in which the two groups would be assumed to be of equal size. For application to any other population, they are therefore weighted by the respective base rates. Let the brain-damaged base rate be π_1, and the non-brain-damaged base rate be π_2 (where $\pi_1 + \pi_2 = 1$). Then the weighted percentage of the brain-damaged group which obtains score x is $\pi_1 p_1$ and the weighted percentage of the non-brain-damaged group which obtains score x is $\pi_2 p_2$. The probability that a person with score x is brain-damaged is given by:

$$P = \pi_1 p_1 / (\pi_1 p_1 + \pi_2 p_2).$$

In Table 2, P has been calculated for all the possible scores of the Modified Word Learning Test. The unweighted percentages obtaining each score have been obtained by combining the data of Walton & Black (1957) and Walton *et al.* (1959,

Table 2. *Modified Word Learning Test: brain damage probability and test efficiency calculated as described in the text*

Raw score	Brain damage probability	Test efficiency
1	—	—
2	1·00	1·00
3	1·00	1·00
4	1·00	1·00
5	1·00	1·00
6	0·35	0·54
7	0·06	0·88
8	0·08	0·86
9	0·03	0·95
10	0·02	0·95

excluding subnormals and ESN). The base rates are those for Netherne described above, viz. $\pi_1 = 0{\cdot}29$ and $\pi_2 = 0{\cdot}71$.

The efficiency of the test used in this way is derived as follows: again only the examinees obtaining the given score are considered. A pronouncement that the

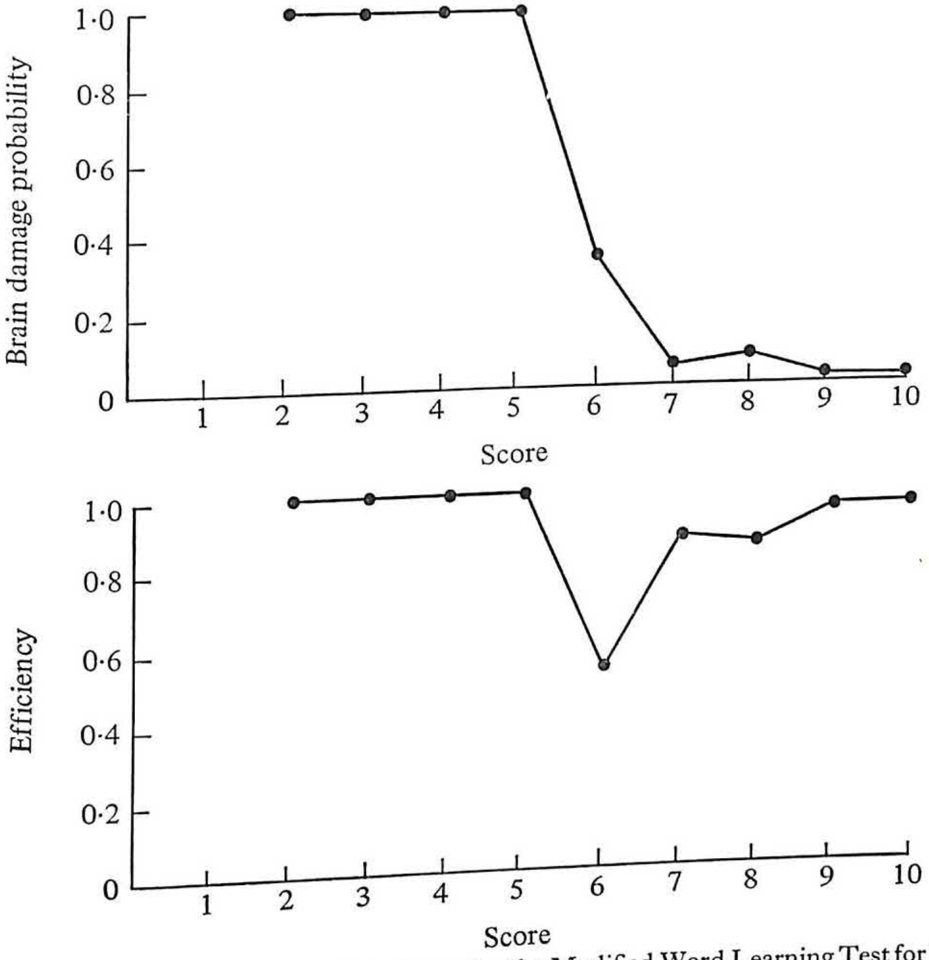


Fig. 1. Brain damage probability and efficiency for the Modified Word Learning Test for base rate of 29 per cent brain-damaged.

examinee has a particular probability of being brain-damaged is considered equivalent to a pronouncement that the same fraction of all subjects (in both groups) obtaining that score are brain-damaged. Those in the brain-damaged group thus classified as being brain-damaged would be correctly classified: similarly those in the non-brain-damaged group thus classified as non-brain-damaged would be correctly classified. All others would be incorrectly classified.

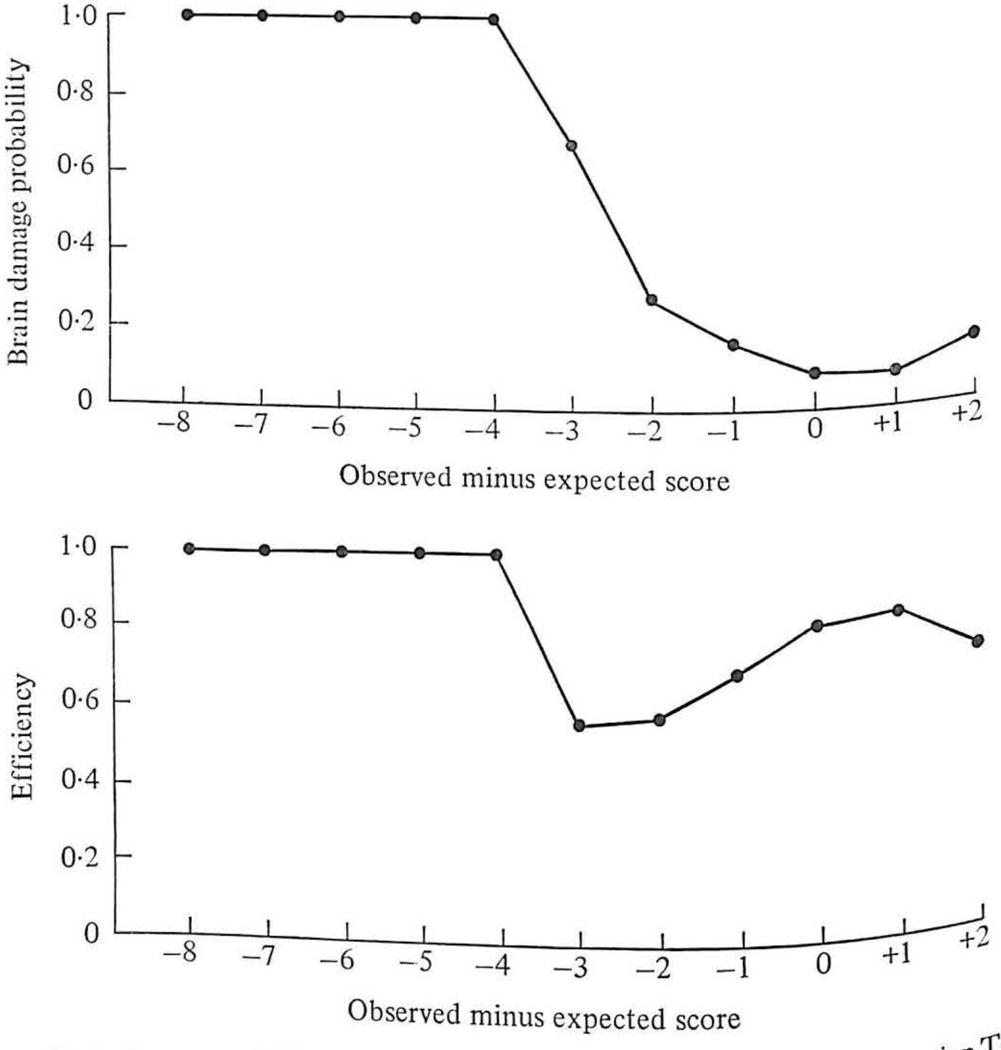


Fig. 2. Brain damage probability and efficiency for the Benton Revised Visual Retention Test for base rate of 29 per cent brain-damaged.

Using the previous notations, the proportion of each group obtaining score x which would be classified as brain-damaged would be P. The weighted percentage of the brain-damaged group correctly classified is then $P\pi_1 p_1$. The proportion of each group obtaining score x which would be classified as not brain-damaged would be $(1-P)$. The weighted percentage of the non-brain-damaged group correctly classified is thus $(1-P)\,\pi_2 p_2$. The total percentage in both groups (obtaining score x) which is correctly classified is $P\pi_1 p_1+(1-P)\pi_2 p_2$. As the percentage in the two groups obtaining score x is $\pi_1 p_1 + \pi_2 p_2$ the total proportion correctly classified—the test efficiency—is given by:

$$E = \frac{P\pi_1 p_1 + (1-P)\pi_2 p_2}{\pi_1 p_1 + \pi_2 p_2}.$$

This has been calculated for the Modified Word Learning Test, using data from the scores given above, and is presented in Table 2.

The brain damage probability and test efficiency given in Table 2 are shown graphically in Fig. 1.

Further examples of the use of brain damage probability and test efficiency calculated in this way are given for the Revised Visual Retention Test (Benton, 1955) in Fig. 2 and for the Memory-for-Designs Test (Graham & Kendall, 1960) in Fig. 3.

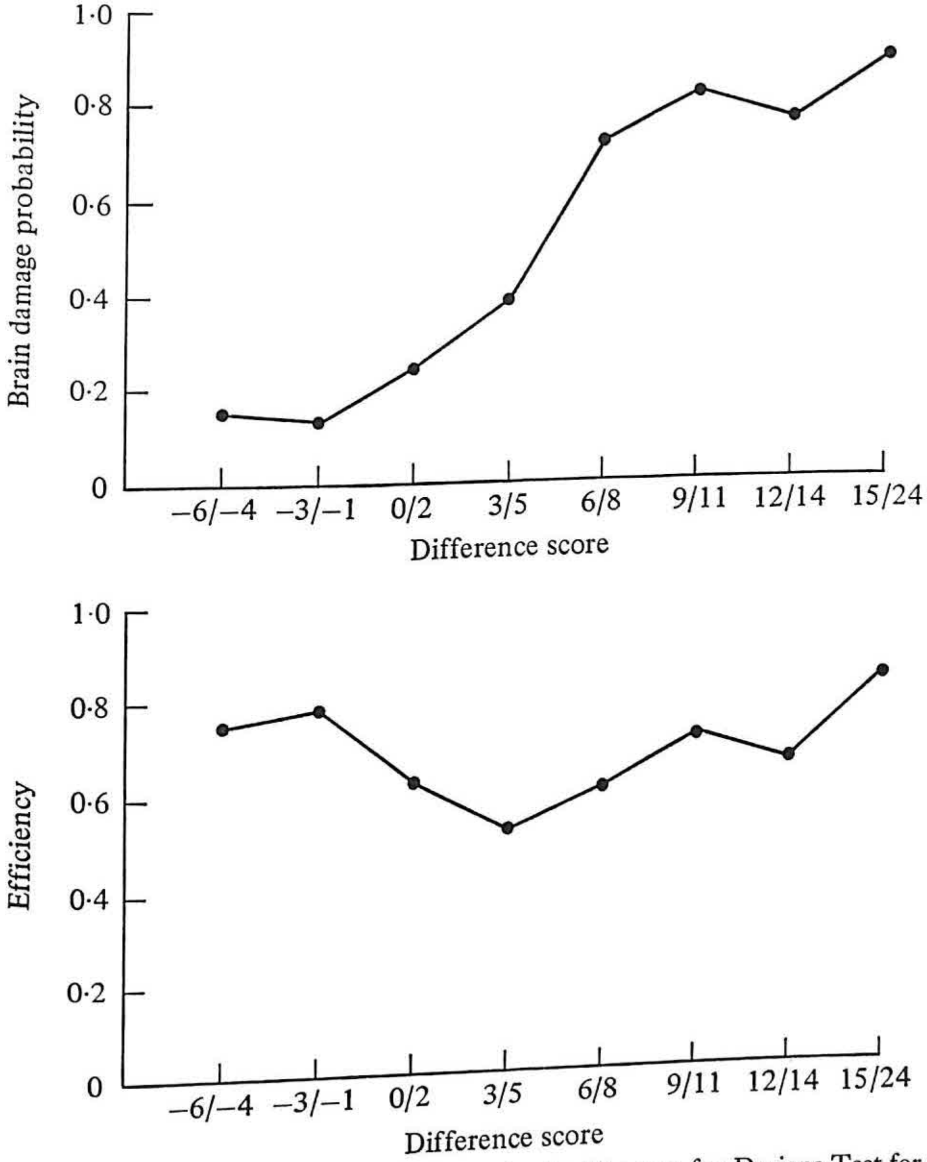


Fig. 3. Brain damage probability and efficiency for the Memory-for-Designs Test for base rate of 29 per cent brain-damaged.

DISCUSSION

The validity of tests for discrimination between normal and abnormal groups is a function, *inter alia*, of the score obtained by the examinee, the base rate of the abnormality, and the method used for calculating the significance of the score. Test scoring methods frequently take account of the score obtained by the examinee: such is the case with the Modified Word Learning Test where it is implicit in the arguments used by the authors to justify the use of a particular cutting score (Walton & Black, 1957; Walton *et al.*, 1959). But the dependence of validity on the base rate and the method used for calculating the significance of the score have usually been ignored. They are not mentioned by, for example, Walton & Black (1957), Graham & Kendall (1960) or Benton (1955), although more recent studies have sometimes included a consideration of the base rate (see, for example, Satz, 1966; Turland & Steinhard, 1969). The test efficiency calculated here is a measure of validity in which dependence on score and base rate has been taken into account. The validity of other methods of calculating brain damage probability has not been considered.

Alternative methods of calculating brain damage probability include: (1) Comparing the numbers of subjects in the two groups obtaining score x or scores further in the brain-damaged direction (Meehl & Rosen 1955). (2) Confining attention to the non-brain-damaged group and considering the percentage obtaining score x or scores further in the brain-damaged direction. (3) Making an actuarial prediction, dependent solely on the base rate (Meehl & Rosen, 1955).

Each of these methods may have the highest validity for some scores on some tests. However, each has some general disadvantage compared with the ordinate comparison method. Methods 1 and 2 have efficiency functions, calculated similarly to that for the ordinate comparison method, which have asymptotes of 1 at the abnormal extreme of the distributions (where they do not overlap) and of less than 1 at the normal extreme of the distributions (where they do not overlap). The ordinate comparison method, however, has complete efficiency for all scores in which the distributions do not overlap. Method 3 has efficiency less than 1 for all scores on all tests.

When the ordinate comparison method is used, the calculated probability of the examinee being brain-damaged (or not brain-damaged) is always less than 1 for all scores in which the distributions overlap. It is for these scores that the method has less than complete efficiency. This relationship of brain damage probability and efficiency of prediction can be seen in Fig. 1.

In addition to the dependence of validity on score, base rate and method of calculation, it is dependent on the extent to which the score distributions and base rates used represent the populations to which they apply. The confidence that can be placed on the base rate used in the example was discussed above. Similar considerations would apply whenever the base rate is ascertained in the same way. Greater reliability would accrue if the base rate were ascertained on successive years and an estimate of the reliability could be obtained by comparing the base rates obtained on two or more successive years. The extent to which the same distributions in the brain-damaged and non-brain-damaged groups represent the score distributions in the populations from which they were drawn can be estimated in two ways.

Firstly, different standardization samples can be compared. The more similar they are, the more confidence can be placed in their representation of the population from which they were drawn. For the ordinate comparison calculations the relevant measures are the proportions in the two groups (brain-damaged and non-brain-damaged) obtaining each score in different standardization samples. This has been calculated for the two samples to provide the norms used in the above example (Walton & Black, 1957; Walton *et al.*, 1959, excluding subnormals and ESN), the results being tabulated in Table 3. It can be seen that none of the score samples differ significantly.

Table 3. *The probability of a difference between the proportions in the two groups: Walton & Black* (1957) *compared with Walton et al.* (1959)

Score	Test	Total n	Probability
1	—	0	—
2	χ^2 (2×1 as adjacent cells contain no entry)	63	$P > 0.10$
3		22	$P > 0.99$
4		18	$P > 0.05$
5	Binomial (adjacent cells contain no entry)	6	$P = 0.11$
6	Fisher exact (2×2)	17	$P > 0.10$
7	Fisher exact (2×2)	20	$P > 0.10$
8	χ^2 (2×2)	50	$P > 0.30$
9		103	$P > 0.50$
10		218	$P > 0.05$

The second method of estimating the representativeness of the samples concerns the continuity of the calculated brain damage probability functions. It is reasonable to assume that this is a continuous function in the population: if the sample gives rise to a continuous function, it is likely that it adequately represents the population in each score category. An arbitrary fluctuation of the function would suggest that the sample is unrepresentative for the score concerned. Figs. 1–3 show continuous functions without arbitrary fluctuations.

If these methods of estimating representativeness show the samples to be unrepresentative, then score categories can be combined until they are. Meehl & Rosen (1955) and Maxwell (1961) contend that very large samples—apparently larger than those normally included in normative samples for brain damage tests—are required for ordinate comparisons as calculated here to be reliable. The high probabilities of Table 3 and the continuous functions of Figs. 1–3 show, however, that they can be reliable in this respect.

I am indebted to Dr H. R. Beech, Mr F. J. Lillie and Miss M. Vaughan for advice and criticism during the preparation of this paper, and to Mr D. Gwyther for assistance with calculations.

REFERENCES

BAYES, T. (1958). Essay towards solving a problem in the doctrine of chances. *Biometrika* **45**, 296–315.

BECK, A. T. (1962). The reliability of psychiatric diagnosis: a critique of systematic studies. *Am. J. Psychiat.* **119**, 210–216.

BENTON, A. L. (1955). *The Revised Visual Retention Test: Clinical and Experimental Application.* Ames: State University of Iowa.

GRAHAM, F. K. & KENDALL, B. S. (1960). The Memory-for-Designs Test: revised general manual. *Percept. mot. Skills* **11**, 147–188.

KENDRICK, D. C. (1965). Speed and learning in the diagnosis of diffuse brain damage in elderly subjects: a Bayesian statistical approach. *Br. J. soc. clin. Psychol.* **4**, 141–148.

KREITMAN, N. (1961). The reliability of psychiatric diagnosis. *J. ment. Sci.* **107**, 876–886.

MAXWELL, A. E. (1961). *Analysing Qualitative Data.* London: Methuen.

MEEHL, P. E. & ROSEN, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns or cutting scores. *Psychol. Bull.* **52**, 194–216.

SATZ, P. (1966). A block rotation task: the application of multivariate and decision theory analysis for the prediction of organic brain disorder. *Psychol. Monogr.* **80**, no. 21. (Whole no. 629.)

TURLAND, D. N. & STEINHARD, M. (1969). The efficiency of the Memory-for-Designs Test. *Br. J. soc. clin. Psychol.* **8**, 44–49.

WALTON, D. & BLACK, D. A. (1957). The validity of a psychological test of brain-damage. *Br. J. med. Psychol.* **30**, 270–279.

WALTON, D., WHITE, J. G., BLACK, D. A. & YOUNG, A. J. (1959). The Modified Word-Learning Test: a cross-validation study. *Br. J. med. Psychol.* **32**, 213–220.

Manuscript received 14 *August* 1970

Revised manuscript received 27 *May* 1971

Br. J. soc. clin. Psychol. (1971), **10**, pp. 375–378
Printed in Great Britain

Hostility and Personality in a Student-Teacher Population

By P. R. MAYO

University of Sussex

AND J. M. BELL

Brighton College of Education

Eighty men and 72 women student teachers completed the HDHQ and EPI questionnaires. A principal component analysis gave support to the HDHQ as measuring hostility, which may be expressed both intropunitively and extrapunitively. Mean group scores in the present study are compared with previously reported studies and found to be higher than expected. The significance of this and the relation of the HDHQ to the EPI measures are discussed.

The Hostility and Direction of Hostility Questionnaire (HDHQ) was designed to measure aspects of both intropunitiveness and extrapunitiveness, and has been applied predominantly to clinical groups (Caine *et al.*, 1967; Philip, 1968, 1969). Norms for groups of normals have been published, but to date there has been no investigation explicitly concentrating on a non-psychiatric population.

In giving the HDHQ to groups of male and female student teachers in training, the opportunity existed to establish norms for a clearly defined age-group, to examine the component structure of the HDHQ for these groups, and to relate the hostility scores to extraversion and neuroticism (Eysenck & Eysenck, 1964).

METHOD

Subjects and procedure. The subjects were 80 men (mean age 21·28 years; S.D. 3·06) and 72 women (mean age 20·46 years; S.D. 2·76 years), all of whom were first or second year students in a college of education. It was not possible to obtain a random sample of the student-teacher population but the subjects were all those who could be contacted by a member of the teaching staff over a six-week period. There was only one refusal. The same person administered questionnaires to the subjects in groups ranging from 15 to 30. Opportunity was given to raise queries both before and whilst they were being filled in.

Measures. The Hostility and Direction of Hostility Questionnaire (HDHQ) (Caine *et al.*, 1967) gives a measure of total hostility (sometimes called general punitiveness) which is the sum of the extrapunitive and intropunitive scales (E+I). The extrapunitive scale is broken down into three subscales: urge to act out hostility (AH), criticism of others (CO), and projected delusional hostility (PH). The intropunitive subscales are self-criticism (SC) and guilt (G). The formula given in the manual for direction of hostility (I–E) is (2SC+G)−(AH+CO+PH), but Philip (1969) has suggested separate scoring for the extrapunitiveness and intropunitiveness scales. This additional scoring is included in the present study.

Form A of the Eysenck Personality Inventory (Eysenck & Eysenck, 1964) was used to obtain E, N and L scores.

RESULTS

A principal component analysis of the HDHQ for both groups gave two main components in each case.

Table 1. *Loadings of HDHQ subscales on the first two components*

	Men ($n=80$)		Women ($n=72$)	
	Component I	Component II	Component I	Component II
AH	0·385	0·471	0·237	0·560
CO	0·387	0·596	0·549	0·266
PH	0·438	0·081	0·424	0·346
SC	0·471	−0·530	0·376	−0·615
G	0·536	−0·369	0·567	−0·343
Eigenvalue	2·022	1·136	1·782	1·485
Variance (%)	40·44	22·72	35·63	29·70

For both men and women the first component involves all the subscales and the second contrasts the three extrapunitive with the two intropunitive scales. Only PH in the men's group is low on component II. Coefficients of similarity between the groups (Barlow & Burt, 1954) were 0·970 for the first component and 0·903 for the second.

The mean group scores for the HDHQ and EPI measures are given in Table 2.

Table 2. *HDHQ and EPI scores for men and women student-teachers*

	Men ($n=80$)		Women ($n=72$)			
	Mean	S.D.	Mean	S.D.	t	P
E + I (Total Hostility)	17·70	6·17	18·72	5·24	1·08	n.s.
I − E (Direction of Hostility)	−2·17	6·41	2·68	7·11	4·39	< 0·001
AH	5·54	2·72	4·92	2·45	1·46	n.s.
CO	5·61	2·53	5·07	1·77	1·50	n.s.
PH	0·90	1·03	0·79	0·98	0·66	n.s.
SC	4·10	2·29	5·60	2·54	3·80	< 0·001
G	1·55	1·38	2·40	1·51	3·60	< 0·001
AH + CO + PH (Extrapunitiveness)	12·05	4·48	10·83	3·85	1·78	n.s.
SC + G	5·66	3·24	8·00	3·58	4·20	< 0·001
EPI E	13·60	4·28	13·14	4·37	0·65	n.s.
N	9·44	4·51	11·67	4·27	3·10	< 0·01
L	2·09	1·65	2·31	1·56	0·85	n.s.

As a group the men are predominantly extrapunitive and the women intropunitive. When the subscales of the HDHQ are considered independently, the women report significantly higher self-criticism and feelings of guilt. The women's group is also higher on N.

In Table 3 product moment correlations are given relating the HDHQ to the EPI measures, and the pattern of correlations is noticeably similar for both groups. Total hostility correlates positively with N and negatively with the L scale but shows no relation with E. Intropunitiveness is negatively associated with E and positively with N. E and N are negatively correlated but this does not reach a statistically significant level in the women's group.

Table 3. *Product moment correlations between HDHQ and EPI measures for men and women*

Men ($n=80$)

		I – E	E	N	L
HDHQ	E + I	0·01	0·06	0·48**	– 0·36**
	I – E		– 0·45**	0·31**	0·05
EPI	E			– 0·33**	– 0·12
	N				– 0·18

Women ($n=72$)

		I – E	E	N	L
HDHQ	E + I	0·14	– 0·05	0·55**	– 0·33**
	I – E		– 0·51**	0·40**	0·12
EPI	E			– 0·22	0·01
	N				– 0·13

** $P<0\cdot 01$.

DISCUSSION

The first two components obtained from the analysis of the HDHQ do support the authors' claims to be measuring a unitary phenomenon which they label hostility, and that hostility may express itself in an intropunitive or in an extrapunitive direction. The PH subscale for men on the second component is the one exception.

The total hostility scores reported in Table 2 are significantly higher ($P<0\cdot 01$ for men, $P<0\cdot 001$ for women) than the figures quoted by Philip (1968) for 240 normals, and in fact are very similar to the means for the neurotic groups reported in that paper. However, the N scores for men are close to the norms given by Eysenck & Eysenck (1964) for the normal population, and those for the women, whilst being somewhat above the normal population scores, do not reach the level given in the manual for neurotic groups. Two possibilities suggest themselves here. It could be that student teachers have higher total hostility than some other occupational groups. Alternatively total hostility may be negatively associated with age in adults. If the latter is shown to be the case it remains to be shown whether young adults are merely more willing to admit to extrapunitive and intropunitive tendencies or whether with increasing age people do in fact become less extrapunitive and/or less intropunitive. A longitudinal study is required to disentangle such factors.

The women's group scored significantly higher on the self-criticism and guilt subscales than the men (Table 2), and this is closely linked to their raised N scores (SC and N: $r=0\cdot 60$, $P<0\cdot 001$; G and N: $r=0\cdot 50$, $P<0\cdot 001$). Previous studies using the Eysencks' scales have given conflicting results in this respect. Thus Ojha *et al.* (1966), using the MPI with university students, did find significantly higher N scores in women students when compared with men students. Ryle & Lunghi (1968), using the EPI also with university students, did not find this difference. Whether women students are higher on these measures when they start their college course or whether they represent a change taking place during the course is not clear at the present time.

The finding that extrapunitiveness is associated with extraversion and intropunitiveness with neuroticism is in line with previous studies (Foulds, 1965; Caine, 1970).

One interesting feature of the results in Table 3 is the negative correlation between total hostility and the L scale. As a consequence of this finding the subjects in each group with L scores of five and over (seven in each group) were omitted and the data reanalysed. The differences from the figures reported in all three tables were very small indeed and clearly non-significant. The possibility that some subjects are 'faking good' (Eysenck & Eysenck, 1964) cannot be excluded, but such an interpretation has been questioned, e.g. Knowles & Kreitman (1965) with adults and Orwin (1971) with children.

REFERENCES

BARLOW, J. A. & BURT, C. (1954). The identification of factors from different experiments. *Br. J. statist. Psychol.* **7**, 52–56.

CAINE, T. M. (1970). Personality and illness. In P. Mittler (ed.), *The Psychological Assessment of Mental and Physical Handicaps*. London: Methuen.

CAINE, T. M., FOULDS, G. A. & HOPE, K. (1967). *Manual of the Hostility and Direction of Hostility Questionnaire*. London: University of London Press.

EYSENCK, H. J. & EYSENCK, S. B. G. (1964). *Manual of the Eysenck Personality Inventory*. London: University of London Press.

FOULDS, G. A. (1965). *Personality and Personal Illness*. London: Tavistock Publications.

KNOWLES, J. B. & KREITMAN, N. (1965). The Eysenck Personality Inventory: some considerations. *Br. J. Psychiat.* **111**, 755–759.

OJHA, A. B., KELVIN, R. P. & LUCAS, C. J. (1966). The use of the Maudsley Personality Inventory on university students. *Br. J. Psychiat.* **112**, 543–548.

ORWIN, W. (1971). A study of the reliability implications of 'lie' scores in the Junior Eysenck Personality Inventory. *Br. J. soc. clin. Psychol.* **10**, 260–266.

PHILIP, A. E. (1969). The development and use of the Hostility and Direction of Hostility Questionnaire. *J. psychosom. Res.* **13**, 283–287.

RYLE, A. & LUNGHI, M. (1968). A psychometric study of academic difficulty and psychiatric illness in students. *Br. J. Psychiat.* **114**, 57–62.

Manuscript received 26 *August* 1970

Revised manuscript received 11 *November* 1970

Br. J. soc. clin. Psychol. (1971), **10**, pp. 379–382
With 1 text-figure
Printed in Great Britain

Cognitive Interference in Spastic-Hemiplegic Children on the Stroop Word-Color Test

By GERALD LEISMAN

Department of Neurology, Manchester University

A research project has been conducted to determine the relationship between efficiency on the Stroop Word-Color Test (decreased cognitive interference) and age in spastic-hemiplegic children. Subjects were presented with three sets of cards: (*a*) consisted of 100 words naming colours, (*b*) consisted of 100 rectangular patches of colour which the child had to name as quickly as possible, and (*c*) consisted of 100 random words naming colours which were printed in an ink of a colour different from the colour named by the word. Cards (*c*) were the measures of the degree of interference. It was found that in spastic-hemiplegic children, unlike normals, the interference effect did not decrease with increasing age. It remained relatively constant.

Although it is difficult to form any general rule when speaking about children suffering from a condition of congenital brain dysfunction, it can be said that there often is a good deal of cognitive interference as a result of such associated handicaps as short attention span, hyperactivity, etc. It is also known that cognitive functioning tends to be more efficient when interference effects from internal and external stimuli are low (Stein & Langer, 1966).

The Word-Color Interference Test (Stroop, 1935) involves the notion that congruent stimuli can be internalized relatively simply and non-congruent stimuli are distracting. The test presents the colour names in non-congruent colours (e.g. 'red' in green ink). The subject is required to name the colours of the ink as quickly as possible and to disregard the written words. Stroop had found that, amongst college students, a longer time is taken to name the colour than the ink. It appears, then, that reading colour words conflicts with the naming of perceptual properties.

Comalli *et al.* (1962) found that in children the degree of interference exhibited on this task is inversely related to the ability to maintain a uniform and directed course of attention when confronted with non-congruent stimuli.

Comalli and his colleagues proposed the hypothesis that development entails an increase in differentiation and hierarchical integration of functions. Efficiency on the Stroop test is based on the ability to differentiate the naming response from the reading response. Therefore one would expect a direct relationship between efficiency on the Stroop test and age, and this has been demonstrated to be the case. The question under study is whether this direct relationship applies to spastic-hemiplegic children as it does to normals.

METHOD

Subjects. One hundred and twenty-seven spastic-hemiplegic children were chosen as they seem to manifest more severe problems to parents and teachers than one would expect from children suffering from a relatively less severe physical handicap than many other brain-damaged youngsters. They also formed a discrete group. The children were attending a centre of recreational facilities for the handicapped, a summer camp for cerebral-palsied children, and various special schools in both the United States and England.

The age range of the children was from 6 to 13 years. They were all of at least average intelligence as rated by the Skemp A.A.C.C. (1968, unpublished), a non-verbal and non-manipulative intelligence test standardized on spastic-hemiplegic children, and Raven's Progressive Matrices (1962). The results were compared with the findings for normal children of the same age range.

Table 1. *Age distribution and mean interference scores for normal and spastic-hemiplegic children*

	Spastic-hemiplegic group ($n=127$)			Normal group ($n=182$)		
Age	n	Mean interference score	S.D.	n	Mean interference score	S.D.
6	8	204	20·3	12	208	16·9
7	14	264	17·1	24	260	18·4
8	17	224	13·2	19	205	8·7
9	16	226	28·6	20	190	6·9
10	20	222	17·3	25	180	13·1
11	22	206	9·8	29	160	17·2
12	16	228	11·2	25	155	16·8
13	14	196	14·1	28	150	7·9

Procedure. The children were presented with three sets of cards containing patches or the names of red, blue, or green. The first set, (*a*), consisted of 100 words naming colours and were printed in black ink and presented in random order. They had to read the words as quickly as possible. The second set, (*b*), consisted of 100 rectangular patches of colour which were also arranged randomly. Again, the colours had to be named as quickly as possible. The third set, (*c*), presented 100 random words naming colours, but the words were printed in an ink whose actual colour was different from the colour named by the word. The task was to name the colour of the ink as quickly as possible. This card was taken as a measure of the degree of interference of the reading of the colour-word on the naming of the actual colours. The measure utilized was the total time taken for responding to the 100 items on each card.

RESULTS

Fig. 1 illustrates the mean total time in seconds for each of the three sets of cards in both the normal and spastic-hemiplegic groups of children. As found in other studies, it appears that a longer time is necessary to name colours alone as compared to naming colour-words. The spastic-hemiplegic children were rated on the Gates–MacGinitie Reading Tests (Primary Forms A, B, C, and Survey D and E tests) (1965). No significant between-group differences were found in their vocabulary and comprehension or speed of reading. The time taken to name colours and to read colour-words is greatest for the youngest age-group and appears to be inversely related to age in normal children. This does not appear to be the case amongst the spastic-hemiplegic children.

Unlike the findings of Comalli *et al.* (1962) for normal children, the interference effect of colour-words on colour naming does not decrease with increasing age of the spastic-hemiplegic children. It remains relatively constant. Both groups of children commence at the same point but digress at age 8.

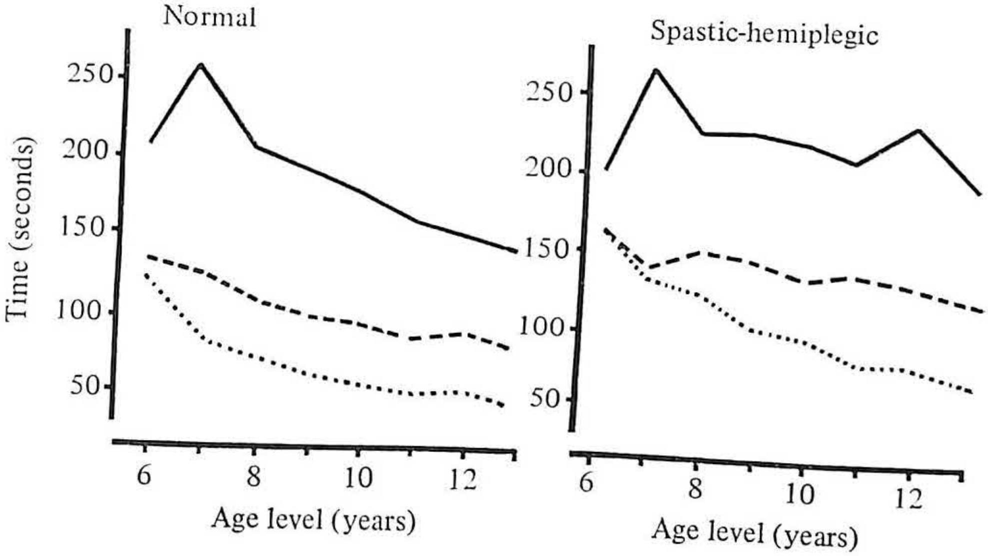


Fig. 1. Interference effects of the Stroop Word-Color Test in normal and spastic-hemiplegic children. ——, colour-word interference; – – –, colours; . . . . ., words.

COMMENTS

It appears, then, that the performance of the spastic-hemiplegic children on the Word-Color Interference Test is inconsistent with Comalli's developmental notion of increasing differentiation and hierarchical integration. It can be assumed that in order to carry out the task adequately, there has to be a differentiation between the colour-naming and word-reading functions. One can also assume that the two operations are of differential potency (i.e. the naming of colours is weaker than the reading of words). Therefore, in order to perform this task adequately, the person must be able to maintain the weaker organization against intrusions leading to shifts to the stronger organizational responses. An increasing capacity to organize such kinds of material in accordance with instructions will then be reflected in a more rapid performance on the interference cards. The results have shown that children characterized by lesser differentiation and little hierarchical integration of functions (i.e. the younger normal child and the spastic-hemiplegic child) perform poorly on interference tasks.

REFERENCES

COMALLI, P. E., WAPNER, S. & WERNER, H. (1962). Interference effects of the Stroop Color-Word Test in childhood, adulthood, and aging. *J. genet. Psychol.* **100**, 47–53.

GATES, A. I. & MACGINITIE, W. G. (1965). *Gates-MacGinitie Reading Tests.* (Primary Forms A, B, C, and Survey D and E tests.) Los Angeles: Western Psychological Services.

RAVEN, J. C. (1962). *Coloured Progressive Matrices.* (Sets A, Ab, B, C, D and E.) London: H. K. Lewis.

STEIN, K. B. & LANGER, J. (1966). The relation of covert cognitive interference in the Color-Phonetic Symbol Test to personality characteristics and adjustment. *J. Personality* **34**, 241–251.

STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *J. exp. Psychol.* **18**, 643–662.

Manuscript received 27 *August* 1970

Revised manuscript received 10 *March* 1971

Br. J. soc. clin. Psychol. (1971), **10**, pp. 383–384
Printed in Great Britain

Simulation of Liberal and Conservative Attitudes on the Conservatism Scale

By RICHARD H. MIKESELL*

Institute for Behavioral Research, University of Georgia

AND J. J. PERSENSKY

University of Cincinnati

Several scales are currently used to assess a variable which has been variously labelled authoritarianism, dogmatism, fascism and anti-scientific attitude (Adorno *et al.*, 1950; Cattell *et al.*, 1957; Eysenck, 1957; McClosky, 1958). These scales require the individual to record some amount of agreement or disagreement to a detailed and relatively lengthy statement of opinion.

The Conservatism Scale was developed as a 'new measure of conservatism' (Wilson & Patterson, 1968). Briefly, the scale is a list of 50 brief labels or catch-phrases which represent familiar and controversial issues (e.g. death penalty, computer music, etc.). Respondents are asked to answer 'yes', 'no', or 'don't know' for each item. A measure of the conservatism–liberalism dimension is obtained by comparing the responses with the scoring key which was devised from three successive item analyses of the scale.

The scale's authors claim that the Conservatism Scale is an improvement on existing scales. Wilson & Patterson reasoned that a two-stage sequential process occurs when individuals answer the traditional items presented in statement form: first, an immediate emotional response, and second, a cognitive response that a rational judgement is expected because the statement is so carefully worded. The Conservatism Scale is an attempt to 'tap' the first stage in the process. Wilson & Patterson contend that items presented as propositional statements tend to increase the effect of social desirability because the long statement facilitates inhibition and consequently cognitive as opposed to affective factors operate when individuals respond to the traditional scale items. Cloud & Vaughn (1970) have demonstrated that the dimension of conservatism measured by the Conservatism Scale is essentially acquiescence-free.

Considerable evidence is available that relatively naive individuals can fake or respond in the direction that the experimenter requests on a variety of measures. For example, Mikesell & Calhoun (1969) demonstrated that subjects can answer so as to appear depressed on a depression inventory. Subjects have also been found to be able to respond so as to appear to be of high self-esteem on a measure of self-concept (Mikesell *et al.*, 1970).

The purpose of the present study was to determine whether or not subjects could simulate liberal and conservative responses on the Conservatism Scale. Specifically, it was hypothesized that the same subjects could significantly simulate both liberal and conservative responses. Sex differences were also examined, although no specific predictions were made regarding this factor.

Subjects were 36 college students, of which 21 were males and 15 were females. Each subject took the Conservatism Scale under each of three instructional sets: I (control set) requested the subjects to 'answer the items truthfully'; II (liberal set) asked the subjects to 'answer so as to appear liberal'; and III (conservative set) instructed the subjects to 'answer so as to appear conservative'. Order of instructional set was completely counterbalanced for the entire sample.

A 2 × 3 (sex × instructional set) unweighted means analysis of variance for unequal n's with repeated measures (Winer, 1962) indicated a significant effect due to instructional set ($F = 24{\cdot}14$; d.f. = 2, 68; $P < 0{\cdot}01$) and a significant sex × instructional set interaction ($F = 33{\cdot}82$;

---

* Requests for reprints should be sent to Richard H. Mikesell, Department of Psychology, University of Georgia, Athens, Georgia 30601, U.S.A.

d.f. = 2, 68; $P < 0{\cdot}01$). The effect of sex was not significant ($F < 1{\cdot}00$). The means are shown in Table 1.

The interaction was probed by the Newman–Keuls technique (Winer, 1962). Examining the sex factor at each of the three different instructional sets indicated that each instructional set significantly differed from every other instructional set ($P < 0{\cdot}05$), except for the control and liberal instructional set comparison for males.

Table 1. *Mean Conservatism Scale scores for the three instructional sets*

	Instructional sets			
	Liberal	Control	Conservative	Mean
Males	37·00	42·00	62·47	47·16
Females	17·52	44·76	73·86	45·38
Mean	25·91	43·61	69·11	46·21

The significant interaction seems to be inconsistent with previous research in which instructional set was manipulated and the sex × instructional set interaction was found to be non-significant (Mikesell & Calhoun, 1969; Mikesell *et al.*, 1970). However, Wilson & Patterson (1968) reported that data from 360 subjects indicated a very small and statistically non-significant sex difference (women score higher and hence more conservatively than men). Data from the present investigation corroborate that particular finding of Wilson & Patterson. Specifically, in the control condition the mean was 42·00 for males and 44·76 for females. The significant sex × instructional set interaction in the present study may have partially been a function of this sex difference inherent in the Conservatism Scale.

More important, however, is the finding that subjects can significantly simulate both liberal and conservative responses on the Conservatism Scale. This finding would seem to have two implications. First, if subjects were so inclined (or motivated), they could successfully produce either a liberal or conservative score on the test. However, under normal conditions of administration, dissembling on an attitude inventory is not generally a problem. Second, the results indicate some positive face validity for the instrument, i.e. if subjects respond to instructions to answer so as to appear liberal (or conservative), they do indeed receive a high liberal (or conservative) score.

REFERENCES

ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J. & SANFORD, R. N. (1950). *The Authoritarian Personality*. New York: Harper.

CATTELL, R. B., SAUNDERS, D. R. & STICE, G. (1957). *Handbook for the Sixteen Personality Factor Questionnaire*. Champaign, Ill.: IPAT.

CLOUD, J. & VAUGHN, G. M. (1970). Using balanced scales to control acquiescence. *Sociometry* **33**, 193–202.

EYSENCK, H. J. (1957). *Sense and Nonsense in Psychology*. Harmondsworth: Penguin Books.

MCCLOSKY, H. (1958). Conservatism and personality. *Am. Polit. Sci. Rev.* **52**, 27–45.

MIKESELL, R. H. & CALHOUN, L. G. (1969). Faking on the Zung Self-rating Depression Scale. *Psychol. Rep.* **25**, 173–174.

MIKESELL, R. H., CALHOUN, L. G. & LOTTMAN, T. J. (1970). Instructional set and the Coopersmith Self-esteem Inventory. *Psychol. Rep.* **26**, 317–318.

WILSON, G. D. & PATTERSON, J. R. (1968). A new measure of conservatism. *Br. J. soc. clin. Psychol.* **7**, 263–269.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

Manuscript received 9 *September* 1970

Revised manuscript received 15 *March* 1971

Br. J. soc. clin. Psychol. (1971), **10**, pp. 385–386
Printed in Great Britain

Positive Affective Stimulus Value and Accuracy of Role Perception

By FREDRICK KOENIG

Tulane University, New Orleans

The idea of examining the way people perceive other persons they know was a diagnostic device suggested by Kelly (1955), and to implement this procedure he developed the Role Repertory Test. While the original Role Repertory Test was useful in generating research in areas of personality and person perception, it was designed basically as a test for clinical use and was unwieldy and time-consuming for most research. A modification of this test was devised (Tripodi & Bieri, 1963) in which subjects would rate stimulus persons on a series of bipolar scales. The stimulus persons are suggested to the subject by such designations as 'a friend you admire of the same sex as yourself', or 'a person of the opposite sex you find hard to like'. In the various versions of this test, half the stimulus persons provided are individuals of positive affect and half are individuals of negative affect. In conducting research on person perception, this researcher noted a tendency for subjects to have a wider range of differentiation for rating negative persons than for positive persons. Simultaneously, Irwin *et al.* (1967) found that cognitive complexity scores were greater for negative persons than for positive persons. They account for this by a *vigilance* explanation that holds that one needs to be astute and sensitive in predicting the behaviour of individuals who are deemed threatening in order to protect oneself adequately. This is a reasonable explanation, but, as the authors suggest, other interpretations are in order. The research reported in this paper is in terms of this author's *justification* explanation. The idea is essentially that there is a trait in our culture, and perhaps others as well, which stresses being friendly, and liking other people. On the other hand, if one does not like someone, one is expected to have a reason for it. To say that someone is 'a nice guy' usually suffices, and that can be the general overall connotation. (It is interesting that a negative counterpart of 'halo effect' does not appear prominently in the literature.) To say that someone is 'a bad guy' usually requires justification. This may develop into a perceptual style of making more differentiation and finer discrimination of negative persons. If it is indeed less commonly expected, and therefore more difficult, to discriminate among people who are thought of favourably, it would seem that those who do so are more precise, sensitive and skilled perceivers of persons. The purpose of this study is to examine the preceding statement empirically, the hypothesis being: the degree to which a subject differentiates persons with positive affect will be directly related to his accuracy of role perception. A modification of the Role Repertory Test was employed (Tripodi & Bieri, 1963) and differentiation was measured by obtaining the numerical difference between the object person with the highest score and the object person with the lowest score for each subject. The difference scores were the measures of differentiation. This was done for the four positive and four negative persons in the test. The accuracy of role perception was based on the Sarbin Stick Figure Test (Sarbin & Hardyck, 1955). 'Correct' modal responses describing the stick figures were obtained from 220 undergraduate judges. The six figures that showed the highest consensus were item analysed and three figures were selected. They depicted 'modelling', 'leaning' and 'sneaking'. The number of correct responses comprised a scale of 0–3. These two tests were presented to 83 undergraduate sociology students. (As in previous studies, there was a sizable difference in differentiation. The medium range of evaluation was 6·3 for positive persons and 9·7 for negative persons.) The range of evaluation scores were divided, as close as possible, at the median into high and low categories. This was done for scores judging both positive and negative persons. The scores on the Stick Figure Test were also divided at the approximate median into high and low categories. These dichotomies were arranged into two 2×2 contingency tables (see Table 1). Using the chi-square measurement of significance and Yule's coefficient of association, Q, one can compare the two sets of results. The degree of differentiation for positive persons is related to accuracy of role perception at the 0·01 level of significance, and the measurement of association is 0·56. The relationship between the degree of differentiation of negative persons and accuracy of role perception, as shown by the table, is practically non-existent. Although the *vigilance* theory cannot be completely ruled out, the implied function of finer discrimination in vigilance is accuracy for predicting undesirable behaviour. Justification merely requires the appearance of 'being fair' in judging negative persons, but not necessarily accurate. The results indicate that differentiation of positive persons is related to accuracy of role perception as measured by stick figure identification, while differentiation of negative persons shows no relationship to accuracy of role perception.

Table 1. *Degree of differentiation of persons of positive and negative affect and accuracy of role perception*

Accuracy of role perception	Positive differentiation		Negative differentiation		Total
	High group	Low group	High group	Low group	
High group	26	13	17	22	39
Low group	16	28	21	23	44
Total	42	41	38	45	83
	$\chi^2=7{\cdot}60$, $Q=0{\cdot}56$, $P<0{\cdot}01$		$\chi^2=0{\cdot}14$, $Q=0{\cdot}08$, n.s.		

REFERENCES

IRWIN, M., TRIPODI, T. & BIERI, J. (1967). Affective stimulus value and cognitive complexity. *J. Person. soc. Psychol.* **5**, 444–448.

KELLY, G. A. (1955). *The Psychology of Personal Constructs*, vol. 1. New York: Norton.

SARBIN, T. R. & HARDYCK, C. D. (1955). Conformance in role perception as a personality variable. *J. consult. Psychol.* **19**, 109–111.

TRIPODI, T. & BIERI, J. (1963). Cognitive complexity as a function of own and provided constructs. *Psychol. Rep.* **13**, 26.

Manuscript received 15 *September* 1970

Revised manuscript received 15 *March* 1971

Br. J. soc. clin. Psychol. (1971), **10**, p. 387
Printed in Great Britain

The Personal Illness Continuum

By G. A. FOULDS

Torret 49, *San Luis, Menorca, Spain*

Three statements in a recent paper (Gordon & Gregson, 1970) indicate a misunderstanding of the continuum of personal illness which I suggested might be worth consideration. They are: (1) 'The total scores [on the Symptom-Sign Inventory] ought to indicate the severity of the personality disintegration if it bears the relation to theory which Foulds (1965) outlines'; (2) 'As the SSI relates to this theory, a factor analysis should yield factors of integration and psychoticism' and (3) 'It is apparent that the SSI is the coalesced resultant of at least five underlying variables, the dominant one being paranoia/non-paranoia. The N group [i.e. non-paranoid] should not, if Foulds' theory is correct, have a mean lower than the P group [i.e. paranoid]'.

The continuum which I suggest is of classes and not of groups within classes. The non-integrated psychotic class should be made up of paranoid and non-paranoid schizophrenic groups and the integrated psychotic class of paranoiac, manic and depressive groups.

I have argued that symptoms which are common to the groups within the integrated psychotic class should be affirmed with roughly equal frequency by the two groups within the non-integrated psychotic class. In addition there should be symptoms common to the two groups within the non-integrated psychotic class which are more or less non-existent among the groups within the integrated psychotic class. This leaves many SSI symptoms unaccounted for (the differentiae between groups) and one is not obliged to believe that all of these symptoms will be affirmed more frequently by non-integrated than by integrated psychotic groups, or even than neurotic groups. If this be so, one cannot assume that non-integrated psychotic groups, or the non-integrated psychotic class, should have a higher total SSI score. Sixty-three per cent of paranoid schizophrenics claimed that they were being plotted against through no fault of their own; 15 per cent of non-paranoid schizophrenics made this claim and less than 10 per cent of each of the remaining six groups did so. This symptom is not, therefore, common to all groups within the integrated psychotic class and can, therefore, have no bearing on a theory relating to classes. The total SSI score will be a mixture of between-class and between-group differentiae and symptoms which distinguish each psychiatric group from normals. I see no necessity for predicting the total scores of classes or groups.

Should a factor analysis based on paranoid and non-paranoid schizophrenics yield a psychoticism factor? A psychoticism factor must be constituted by symptoms affirmed with considerable frequency by all groups within the psychosis class. Since not all psychotic and non-psychotic groups (from whom the former should be differentiated) were included in the factor analysis of Gordon & Gregson, one could not expect to find a psychoticism factor.

REFERENCES

Foulds, G. A. (1965). *Personality and Personal Illness.* London: Tavistock Publications.

Gordon, A. V. & Gregson, R. A. M. (1970). The Symptom-Sign Inventory as a diagnostic differentia for paranoid and non-paranoid schizophrenics. *Br. J. soc. clin. Psychol.* **9**, 347–356.

Manuscript received 2 March 1971

Br. J. soc. clin. Psychol. (1971), **10**, p. 388
Printed in Great Britain

A Comment on Foulds on Continua

BY R. A. M. GREGSON AND A. V. GORDON

University of Canterbury, New Zealand

Foulds (1971) has made some welcome comments on an interpretation of his SSI in the context of our paper (Gordon & Gregson, 1970) on its use as a diagnostic differential instrument.

The use of the total SSI, or a weighted subset of it, as a differential instrument does not rest necessarily on its theoretical status; it might function as an effective diagnostic instrument in the way reported whatever the underlying psychological processes are which lead to responses being made. The establishment of a differential instrument is a problem in multivariate sampling theory; it can precede or follow the establishment of a coherent theory concerning why it works, or should work. However, it is more satisfying to relate it to a theory.

Foulds (1965) in his original formulation used set-theoretical terms, and suggested a general continuum of illness severity which was incompletely defined; it is not a strong ordering on his partitioning of classes and groups. Further, the SSI itself is a set of measures which are representable in a vector space of total scores and subscale scores of items weighted 0 or 1, but the rules for translating from the scores to the set theory are not fully worked out. He offered some results on the distinction between paranoid and non-paranoid schizophrenics; these were partially confounded with the paranoiac–paranoia and integrated–non-integrated distinctions, and he suggested that paranoids as a group were not homogeneous in their degree of integration, and that non-integrated paranoids belonged with non-integrated schizophrenics. A sample of non-paranoid schizophrenics should on average be more disintegrated than paranoid schizophrenics, if his results were replicable. Our results do not seem consistent with this so far as our N and P comparison goes. The SSI theory is only weakly structured, in terms of the relative psychometric contribution of the various effects distinguished to the generation of between-class and between-group differences.

It would be helpful if researchers trying to make quantitative sense out of the qualitative complexity of SSI structure could set up predictions for the differential value of each component item. They are forced to an exploratory multivariate reanalysis precisely because the psychometric structure of Foulds' theory is incomplete.

Foulds' remaining comments seem to rest on a confusion between the summation of weighted and unweighted component scores; and leave no room for the possibility that his diagnostic framework of groups and classes may be itself error-loaded, which is somewhat sanguine. A psychoticism factor may be generated by individual differences in the severity of illness of a given type within a class, as well as by between-class differences.

REFERENCES

FOULDS, G. A. (1965). *Personality and Personal Illness*. London: Tavistock Publications.

FOULDS, G. A. (1971). The personal illness continuum. *Br. J. soc. clin. Psychol.* **10**, 387.

GORDON, A. V. & GREGSON, R. A. M. (1970). The Symptom-Sign Inventory as a diagnostic differentia for paranoid and non-paranoid schizophrenics. *Br. J. soc. clin. Psychol.* **9**, 347–356.

Book Reviews

Family Constellation: Its Effect on Personality and Social Behavior. By Walter Toman. New York: Springer, 1969. Second edition. Pp. viii + 280. $5.75.

This is in some ways a tantalizingly disappointing book. The central theme is that an individual's sibling position, together with that of his grandparents, affects his character formation, his relationships and his attitude towards authority, work, religion and so on. Dr Toman presents his thesis in the form of games theories, but unlike the sensitive humour that abounds in Eric Berne's *Games People Play*, the author makes this game heavy with seriousness in an off-putting way. Nevertheless, this is an exciting idea and it seems a pity that the presentation falls between the two stools of lightheartedness and seriously careful analysis. For example, according to Dr Toman's thesis, if the older brother of a sister(s) marries the younger sister of a brother(s), they are getting in marriage precisely the (peer) relationship they had at home. He is used, as an elder brother, to a younger sister, and vice versa; so there are likely to be no seniority conflicts, and both of them are used to the other sex, so they should have no sex conflicts either.

This seems to be a global view of both individual and family relationships. To give Dr Toman his due, he does stress the effect of age differences in sibling positions; for instance, too much age gap tends to lead the child developing as a singleton. He also stresses the importance of losses and adds that in order to speak intelligibly, the second- and third-order determinants have to be ignored.

The persons studied have either been in psychoanalysis or psychotherapy with the author —51 cases, 45 cases for diagnosis only; plus 58 among his friends and acquaintances; plus a further 108 have been seen in counselling or psychotherapy with supervisees of the author. From this material Dr Toman has delineated 64 major types of conflict that can arise between a person's sibling position and that of his parents. The message is that everything is explicable in these terms.

Computers show up similarities, which is presumably what Dr Toman is after; this is valid. However, I regret that there is no allowance for the life cycle of the individual, or for the significance of other important relationships which can and do effect personality change. Dr Toman states quite accurately that new relationships for the individual will be determined by the old, but nowhere does he allow for the unexpected inventiveness which seems to be part of human nature. Nor does he allow for the surprising deteriorations which can sometimes occur, as for example schizophrenic withdrawal.

As this book is aimed widely to professional care-givers, researchers and teachers, I hope that Dr Toman will refine and develop his idea further, but either as a serious scientific piece of research with validating data, or as a delightful game which readers can play, but not take too seriously.

MARGARET ROBINSON

Readings in Adolescent Development. By Harold W. Bernard. Scranton, Penn.: International Textbook Co., 1969. Pp. 482. $2.50.

This book of readings is divided into four parts and nine sections and contains 50 separate contributions. At first glance it appears to be directed primarily at American students, if the content of many of the articles is taken as the criterion. Examples of such articles are 'American Teen-agers of the 1960's—Our Despair or Hope?', 'Children of the American Ghetto' and 'The Split-Level American Family'. In all, 18 of the 50 chapters are heavily biased towards the American scene, whilst a further 33 are more general in nature, leaving one ('Youth and Youth Services in England' by Oettinger) specifically English in orientation.

However, if one's perspective is that of a scholar of comparative education, or if one is interested in studying the problems of adolescence in a culture somewhat dissimilar from our own (albeit becoming less dissimilar each day), then the book has much to offer. Among the contributors are such people as Frank Musgrove, Carl Rogers, Robert Havighurst, J. P. Guilford, Harold Loukes, Erik Erikson, Jerome Bruner, Paul Torrance, Nevitt Sanford, David Ausubel and Urie Bronfenbrenner. For their contributions alone the book is worthwhile, although other less well-known contributors supply important and clearly written articles on problems relevant to our time.

It is always interesting to note how people from abroad perceive and record what they see of our educational system. Oettinger's article on youth services in England is a case in point. Her description of regional, area and county colleges is somewhat foreign to the reviewer, whilst the three problems she heard mentioned most were dead-end jobs for the unskilled, early school-leavers and homosexuality. Her view of the English adolescent scene is of value and interest to those who teach the age-group.

In a book of under 500 pages, and one which includes 50 contributions, there is not much room for long expansive articles. Their average length is only nine pages, with individual contributions ranging from five pages to 14 or more. Each of the nine sections is preceded by a statement of what is to be found in the chapter. The statements are full and contain what Bernard considers to be the main points. Not all of the contributors are employed directly in the field of education, a notable exception being Charles E. Wyzanski, Jr., a Federal judge. Also, not all of the contributions include a bibliography, though several that do give extensive lists of books and articles. There is a useful 89-item glossary, and both a name and a subject index.

The articles cover a fairly broad field not indicated in the title. Guilford, for example, writes on creativity, whilst Carl Rogers discusses 'A Humanistic Conception of Man'. Two of the chapters are specifically concerned with the drug problem, whilst that of Kvaraceus concentrates on alienated youth both in America and abroad, a problem of particular interest at the present time.

Although some of the articles have a restricted application in Britain, the total impact left by the volume is that it has something to offer a wide range of readers interested in the problems of the adolescent.

ALFRED MAYFIELD

Moral Education. By N. J. Bull. London: Routledge & Kegan Paul, 1969. Pp. 183. £1.25.

The essential weakness of Dr Bull's book is that it reads like an abridgement of a longer work. It falls, in a general way, into three sections, of which the first is a discussion of a psychological investigation into the development of moral judgement in children and adolescents, the second deals more generally with the problems underlying moral education, and the third contains advice to teachers and intending teachers. Unfortunately, there is not space in a work of this length for any of these very large topics to be worked out, so that the treatment is lacking in penetration, with awkward changes from one subject to another. This is not helped by the excessive division of chapters under subheadings and sub-subheadings, particularly since these are not always as well chosen as they might be. Thus the first subheading for Chapter 11 contradicts the chapter heading, while the subheading 'developmental learning' occurs twice in successive chapters.

The inadequacy of development is most apparent in the first section, where there is neither a systematic account of the experiments upon which the discussion rests nor an analysis of the results, so that the reader is unable to form any assessment of Dr Bull's conclusions. In the later sections there is a need for the kind of analysis of the concepts underlying moral education which one finds, for example, in the work of R. S. Peters or John Wilson. There is, in fact, much useful and sensible discussion in these sections, notably on the relation between religion and ethics and upon moral education in adolescence, but, in the absence of any conceptual analysis, some of the author's most important contributions lack adequate foundations. For instance, authoritarianism, physical discipline and indoctrination are rejected as instruments of moral education, but the grounds for making the rejection are no more than the authority of the writer. Similarly, there is no mention, or apparent awareness, of the dangers attaching to indirect, or accidental, moral learning, and in particular of its irrational nature.

In general, the book fails through trying to hit two targets at once. The critique of Piaget's stages of moral development will not satisfy the psychologist, while the teacher will look for a fuller and more closely reasoned foundation for his moral teaching.

DAVID WARDLE

Kids, Cops and Kilos: A Study of Contemporary Suburban Youth. By G. Thomas Gitchoff. San Diego, Calif.: Malter-Westerfield Publ. Co., 1969. Pp. viii + 178. $2.95.

Dr Gitchoff's book is paved with good intentions. The first chapter holds the promise of description and analysis of changes among contemporary suburban youth, with particular reference to youthful deviance and the emergence of the 'hang-loose ethic'. In addition he promises to explore the relationship of this piece of research to earlier work on American suburbia, the problem of increasing crime rates among juveniles, and the observed phenomena of youthful disaffection with contemporary society. In the end, the reader is disappointed and confused.

The study concerns Pleasant Hill, an upper middle-class suburb 30 miles from San Francisco. Close association (by means of participant observation) with a number of teenagers results in the author describing the interests, attitudes and activities of these people in 1966, and monitoring changes through to 1968. The major change shown is an increasing disaffection with 'typical' middle-class norms, values and activities, and the emergence and adoption of the cool hang-loose ethic.

A wealth of descriptive material, however, gives little theoretical yield on discussion, either concerning the social processes involved in the observed changes (for example, the process by which drug taking is initiated and maintained) which are largely ignored, or in the underlying causes for the widespread changes. Here the explanation given is in terms of youthful disaffection with contemporary American values and of calling attention to the situation via the mechanism described. By adopting new values youth is searching for a way to 'guide [the country] back to the concerns of humanity, humanism and hope'. Other issues raised in the first chapter are forgotten or assumed.

There is a confusion in the work as the author collected data from two viewpoints: initially, that of a researcher, but for the last eight months, while Director of the Youth Services for the City of Pleasant Hill, involved in an action role. Far from having 'no axes to grind' (p. 13) the author was at times running a programme incorporating some of his observations and explanations. In this situation political and administrative criteria were relevant. The complaint is not that this cannot be done, only that it raises problems of observation and explanation which are ignored in the text, and which leave the reader confused as to the framework within which explanations are being presented. The relationship of research to action is described briefly, but the implications of this process are lost.

JANET FORD

Authority and Addiction. By Leon Brill and Louis Lieberman. Boston: Little, Brown, 1969. Pp. 318. $13.50.

Considering the relatively minor incursion in numerical terms of narcotic dependence into British and American society, and when placed beside the rate of innovation of techniques in the areas of rehabilitation associated with mental illness or imprisonment, the effort and imagination devoted to the treatment of addicts are genuinely remarkable.

Brill and Lieberman describe and evaluate one action research project supported by the current enthusiasm. Carried out by the Washington Heights Rehabilitation Centre and sponsored by both local and national welfare agencies, the project was directed towards male narcotics probationers of all ages, and was based on three distinctive procedures. Firstly on its use of 'rational authority', a system of graduated sanctions ultimately enforceable through the power of the courts, designed as a holding device to sustain the addict through the lapses and crises which appear to have vitiated other rehabilitation enterprises. These sanctions ranged from the personal disapproval of the social worker in charge of the case to a prolonged remand to prison. (In this respect, the Washington Heights programme was markedly different from the Synanon experiments in which informal social influence was the principal sanction.) Secondly, on an aggressive casework strategy labelled 'reaching out' in which the Centre penetrated home, neighbourhood and hospital milieux to initiate and maintain treatment relationships. Thirdly, the direction of each case was placed jointly in the hands of a probation officer and a member of the Centre (a social worker or nurse). This involved a blurring of roles and a pooling of information not conventional among welfare professionals. The programme was assisted by the routine use of thin-layer chromatographic analysis of urine samples, which gave accurate indications of any current drug use by probationers. Great stress was laid throughout on the general social functioning of the addict rather than demanding total abstinence as a condition for treatment.

For the purposes of evaluation, two samples were drawn, group I ($n = 94$) receiving the full Centre service and group II ($n = 86$) acting as a control by being dealt with solely by the Office of Probation. Indices for measuring changes in heroin use, criminal activity, work record and social conventionality were constructed. In the event, both groups showed significant improvements, group I performing very marginally better overall. The investigators conjecture that, had the evaluation taken place after a longer period of rehabilitation (a maximum of a year), the differences would have been more dramatic.

The final chapter is devoted to the presentation of a persuasive conceptual model of the life cycle of addiction. It is argued that the particular stage reached in the addict's career (with its corresponding self-image) and the extent of the social skills developed prior to addiction both require assessment and reflection in any rehabilitation programme. The authors reasonably urge that a differentiated addict population requires the availability of an array of treatment milieux together with the possibility of transfers between them.

Cultural differences notwithstanding, the design of both the programme and its evaluation deserves wide exposure to those in Britain interested in the management of treatment of all those categories of deviants we archly think of as having 'character disorders'.

KENNETH LEVINE

Behavioural Sociology: The Experimental Analysis of Social Process. Edited by R. L. Burgess and D. Bushell, Jr. New York and London: Columbia University Press, 1969. Pp. xiii+418. $12.50.

Robert Ardrey has written that 'the primacy of the conditioned reflex contributed lasting damage to American psychology and, like some great wave which gathers fury and greater destructive power the further it moves from the storm centre, brought to other departments of American learning devastation typical of a disaster area'.* If Ardrey's perspective is correct, then the title of this volume of papers suggests that it needs to be considered as the first direct onslaught of the wave upon the structure of sociology.

The editors allege that 'there is a new and growing field within experimental psychology which has a good deal to say to our own discipline, sociology', and the declared purpose of their book is 'to draw together a number of articles describing the experimental analysis of behaviour in terms which might recommend its further consideration by sociologists and social psychologists'.

Measured by its own objectives the book does not appear to be notably successful, for although at some points some of the rigour of the laboratory is transferred with advantage to certain areas of sociological thought, most of the sociological contributors appear either to run into difficulties in applying behavioural psychology to the problems of the sociological world, or to feel the need to express reservations about its application of a kind which negate the underlying thesis of the book... the thesis that a 'behavioural sociology' is about to appear.

Contradictions begin in the prologue by Homans. Homans maintains that 'the principles of behavioural psychology are the general propositions we (i.e. sociologists) use, whether implicitly or explicitly in explaining all social phenomena', and that any failure to accept this stems from a mistaken view of the nature of theory, or from a 'trained incapacity to recognize the part played by psychological propositions in social theory'. And the key point upon which Homans appears to rest his support for this contention is the assertion that 'There is some question whether there is any other general psychology. Those who claim to be dissatisfied with behavioural psychology are hard put to state what propositions they would propose to replace those of behaviourism, or their own propositions can be shown to follow as conclusions under special circumstances from behaviourism'. Thus Homans justifies his sweeping statement that all sociology rests upon behavioural psychology only by equating behavioural psychology with all psychology.

---

* ARDREY, R. (1967). *The Territorial Imperative.* London: Collins.

The closed circle of logic, irrefutable in its own terms, which Homans draws around his position echoes the closed circle which surrounds behaviourism itself: as Janet Daley has recently written, 'the behaviourists' complaint about Freudianism is that it interprets all data in its own terms and as such is never susceptible of disproof *as a system* . . . this is equally the case with behaviourism . . .'.

The body of the book is divided into three parts. Part I provides a clear and painstaking introduction to the principles of behavioural psychology, and proceeds to illustrate these principles with a series of laboratory studies, including some new variations of small-group communications experiments by Burgess. Included is a most striking account by Baer, Peterson and Sherman of a series of experiments which enabled them successfully to shape aspects of the behaviour of severely mentally retarded children, using basic operant principles.

After this helpful introductory section, the second part of the book is intended to carry the experimental approach of the behaviourist out of the laboratory in order to apply it to the analysis of social process in field situations. The experiments are described with meticulous detail, and apart from Wolf, Giles and Hall's apparent unawareness of the possible consequences of the Hawthorne effect, the experimental designs are impeccable.

After these exhaustive preliminaries, then (nearly 300 pages of them), one comes to the third part, the core of the book. At this point, it should be demonstrated to the reader that behavioural psychology really does have some applicability to sociology—but it is at precisely this point that the book is most disappointing.

The most lucid contribution, and the one paper which directs itself most usefully to the main thesis of the book, is the paper by Emerson on 'Operant Psychology and Exchange Theory'. Emerson first makes a point, which the editors have ignored, that operant 'explanations' of social phenomena often amount to little more than translations to a more general level, with the explainer overlooking the losses incurred in such translations. 'In this process', Emerson notes, 'the sociological task is left untouched'.

Emerson moves on to point out that if operant psychology is to be used by the sociologist, the sociologist must first reduce the generality of operant principles by introducing new attributes into them relating to social structural attributes, and then focus his attention upon these added attributes. He goes on: 'If we start with a general proposition and then reduce the level of abstraction of its concepts, the proposition will (i) become less general, that is, gain more specific or restricted empirical focus (toward the subject matter of sociology), and (ii) remain "true" if it was true at the outset. But this "truth" is purely analytic, and it might not show up empirically. Having introduced additional empirical attributes, these might in turn introduce confounding variables. Or conversely, if the less general statement is empirically true, its truth might depend upon these additional attributes rather than upon the more general proposition alone'. And, in the absence of any empirical work which can be discussed, that, in a sense, is all that needs to be said about behavioural sociology. However, there is a number of other issues which are raised by the main thesis of this book which are either not mentioned, or are not dealt with adequately in the book.

It is myopic to deny the importance of various forms of conditioning in obtaining an understanding of certain basic aspects of both animal and human behaviour. To this extent, any sociologist who disregards the conditioned response and its associated phenomena *at those points at which it is relevant for his work*, does so at his peril. The rub comes, of course, in deciding upon the points at which it *is* relevant.

To judge from the contributions to this book, the juxtaposition of behavioural psychology and sociology may help in bringing a more rigorous scrutiny of many of those sociological theories which are loosely or haphazardly conceptualized (as Kunkel's paper shows), and it may also provoke further discussion and a more enlightened understanding of the relation between all forms of psychology and sociology.

But after all this, one must ask whether the operant response is enough. The basic units of experimental behavioural psychology are so small that in order to predict, control, or even try to explain the behaviour of, say, a one-year-old, the existence of many enormously long 'chains of reinforcers' must be postulated. And since the problems faced by sociologists are far more complex than this, containing 'conflicting or ambiguous standards of reinforcement . . . poorly established deprivation characteristics, changing or uncertain discriminative stimuli, inadequately learned behaviour chains, and unknown or shifting reinforcement probabilities' (Kunkel), the sociologist has need of conceptual constructs which can be handled at a higher theoretical level than can those of the behavioural psychologist.

At one point in their discussion, the editors remark that 'Occasionally, however, individuals are observed who do not seem appropriately responsible to (punishing) stimuli' and one hopes that one sees here the beginning of a realization that people and children carry out very complex internal processes of 'figuring out' what their situation means to them and what other people are trying to do to them. Behaviour dependent upon their own complex attributions of meaning to their world is consequently beyond the comprehension of the behaviourist *per se*, and cannot be dealt with by means of his explanatory system. To illustrate this point, one may take the examples of 'token reinforcement' therapy described in Part II, particularly the remedial work carried out by Wolf, Giles and Hall. From the remedial point of view, the experiment was successful, for the children showed a clear improvement in academic performance when compared with a control group. What is not clear is how far these results were achieved solely by the exercise of the principles of behavioural psychology, and how far they were dependent upon the understanding which the students had of the expectations of the experimenters, and of the system which the experimenters had devised for them. The really fundamental psychological and sociological problem which is unanswered in such experiments is that of explaining how the subjects managed to manipulate their own behaviour to produce the results that the experimenters wanted!

In industry, when workers are subjected to a piece-rate system of payment for work, making use, as Skinner points out, of a ratio-schedule of reinforcement, the system works well up to a point; but the degree of understanding and the sophistication of the manipulation with which workers regulate their responses to such payment systems are far more complex than the behaviourist's simple 'stimulus-response' model even begins to suggest.

This brings us to the book's political and ideological implications. Daley* has suggested that behaviourism contains an ideology of manipulation which in practice 'regards the *utility* and *conformity* of the human being as a societal unit as the ultimate and all-embracing value. Evaluating the worth of the demands of a given society *on* individuals is *practically* and morally beside the point; it is the individual who must be made to fit . . .'.

With one exception, none of the papers in this book takes into consideration the moral aspects of the kinds of human manipulation which are being described, or even show any awareness that moral issues might be relevant. Their stance may be described as a 'scientific amorality'. And yet, not only do many of the experiments described 'work' because they degrade and devalue the individual subjects; they also seem in a number of instances to make use of a number of practices which give rise to ethical doubts. Hawkins, Peterson, Schweid and Bigou devised a punitive stimulus for four-year-old brain-damaged Peter, of getting his mother to put him in an empty room for a minimum of five minutes. On three occasions when he was put into isolation during the experiments Peter broke the windows of the room, cutting himself twice. Bushell, Wrobel and Michealis devised the punitive stimulus for a class of three- to six-year-olds of leaving behind in the classroom those children who had not earned enough tokens to purchase 'a special-event ticket', while the rest of the class went off to a theatre, a park or a film show. And so on.

But why is it that we feel only ethically doubtful about these practices, and not morally outraged? Emerson supplies us with the clue, when he states clearly that 'the relevance of operant psychology for sociology is, first and foremost, *social power*'. This leads Emerson directly to the conclusion that Skinner has a power advantage because he can and does exert more control over the pigeon that the pigeon can exert over Skinner. Most of the manipulations described in this book are not morally outrageous because they take place within some power relationship which is accepted within the existing social order, and where forms of manipulation exist anyway. As it is acceptable for Skinner to manipulate his pigeons, so it is acceptable for parents to manipulate their children, for teachers to manipulate their pupils, for social workers to manipulate their delinquents, and for hospital staff to manipulate their mentally subnormal patients.

---

* Daley, J. (1969). A philosophical critique of behaviourism in architectural design. In Broadbent & Ward (eds.), *Design Methods in Architecture*. London: Lund Humphries.

But having realized the centrality of social power to operant psychology, what, then, are we to make of the thesis that a behavioural sociology is on the way? If such a sociology is to be created and used, who defines the problems which are suitable for manipulation in the 'behavioural sociologist society' relationship?

It has been maintained above that behavioural psychology is too simple, and operates at too low a level to be able to achieve the goals set for it by Skinner, or by the editors of this book. But if this view should turn out to be incorrect, and if sociology should become the province of the behavioural psychologist, sociologists may still find that they have a role in devising social means of helping people to resist operant conditioning. For, as Daley says, 'What if the world were really intolerable? Would it do to "recondition" all those sensitive enough to find it so?'

BARRY A. TURNER

Depression: Clinical, Experimental and Theoretical Aspects. By A. T. Beck. London: Staples Press, 1969. Pp. xiv+370. £3.25.

Professor Beck introduces his monograph as 'an attempt to demarcate the body of knowledge regarding depression, to separate what is known from what is alleged or simply taken for granted', and as a presentation of his own theoretical formulation based on the study of over one thousand patients. He discusses the definition of depression, and points to the major unresolved issues regarding its nature, classification and aetiology. The subsequent critical survey of the literature and his original contributions lead back consistently to his views on these basic issues.

Part 1 is devoted to clinical aspects of the problem. First, depression is defined without reference to conventional nosology but in terms of its main attributes, namely a specific alteration in mood, a negative self-concept and self-punitive wishes, vegetative changes and changes in activity level (chap. 1). The symptoms are detailed under four major headings as emotional, cognitive, motivational, and physical and vegetative, with separate headings for delusions (are they not a cognitive disorder?) and hallucinations. As the author himself states, 'some of these divisions may appear arbitrary and . . . some of the symptoms described separately may simply be different facets of the same phenomenon'. The description of the symptomatology, which is clear and vivid, is based on the literature and on the author's own material. Readers may find confusing, however, the fact that while most of the reviewed studies (except Grinker's) deal with depressive illness, Beck's own findings, with which they are compared, relate to depressive syndromes defined independently of primary diagnosis and without regard to nosological categories (chap. 2).

The course and prognosis of depression as a clinical entity is presented in a review of the literature from Kraepelin to contemporary authors (chap. 3). Data on the duration of attacks and intervals, on the likelihood of recurrence, on the relation to age and other factors are quoted, and the information is summarized at the end of the chapter. The reader, being confronted with work performed over a 50-year span, will wonder whether the course and prognosis of depression has been affected by modern methods of treatment and it is regrettable that the author has not tackled this problem.

In chap. 4 Professor Beck deals with the classification of affective disorders. He discusses its reliability and validity (or rather unreliability and lack of validity!) and comments on the controversy that exists between what he terms the unitarian and the separatist schools of thought. The debate on the existence of separate neurotic and psychotic depressive syndromes is summarized in chap. 5. Chaps. 6–8 deal with manic depressive reactions, the controversial topic of involutional melancholia and the problems of schizo-affective reactions.

Part 2 of the monograph is a critical survey of biological and psychological studies of depression. Professor Beck pays careful attention to methodological details, and he emphasizes the controversial nature of the results obtained from biological investigations. He discards most of the reported findings as inconclusive. None of them is regarded as certain, only two receive the label 'probable' (namely, the decreased amount of stage four sleep and the increased secretion of adrenal steroids), and all others are considered as doubtful or uncertain (e.g. data on heredity, glucose metabolism, electrolytes, sedation thresholds, etc.). Concluding this section, Professor Beck writes: 'despite the hundreds of studies in this area, there is little if any solid knowledge of the specific biological substrata of depression'.

With regard to psychological investigations, the author brings to attention a group of findings which suggest that the performance of depressive patients on psychological testing is largely unimpaired and that they tend to underestimate their own achievements. Factor analytical studies are presented as an important new field of exploration, but their results are considered as being tentative only. Studies of personality and of family background, centred round the psychoanalytical studies of Cohen, Beck evaluates as dubious because of methodological inadequacies.

The aim of Professor Beck's own systematic investigation of depression (Part 3) is to find psychodynamic factors specific to the syndrome. The most important aspect of the study is the construction of an inventory for measuring depression and the evaluation of its reliability and validity (chap. 12). This inventory, described in earlier publications, has been widely used by other research workers in the field and its value as a research tool is universally acknowledged.

The initial basic concept, derived from observations on patients with psychotic depression and from the dreams of neurotic depressives, is that a central feature of the syndrome is 'masochism' which is defined as 'a number of distinctive patterns that lead to suffering disproportional or inappropriate to the reality situation'. Testing this hypothesis, the author demonstrates that 'masochistic patterns' (in dreams, early memories, projective tests, etc.) are more frequent in depressive than non-depressive patients and that depressive patients tend to distort their experiences in an idiosyncratic way so as to misinterpret events in terms of failure, rejection and deprivation. The investigations are elegantly devised and the results are interesting, but they do not lead to any new and fundamental concepts. 'Masochism', as used in this context, refers merely to a cluster of symptoms which have long been considered as an intrinsic part of the depressive syndrome. Furthermore, there is an important methodological criticism which may be levelled at Professor Beck's work, namely that it is tautological. The group of depressive patients was delineated by the inventory which itself includes 'masochistic' symptoms, albeit under other names, in several of its items. Thus the conclusions of the study are contained in the initial definition of the material investigated and in the research tool used.

Interesting findings in respect of bereavement are mentioned in chap. 14. While there was no association between childhood bereavement and nosological categories of depression, the group who gave a high depressive score on the inventory had a significantly higher incidence of bereavement prior to the age of 16 than those with a low score.

In Part 4 Professor Beck reviews various theories of depression and discusses neurological, biochemical, psychodynamic, and psychosomatic formulations as well as the existential approach (chap. 16). Concepts of the psychological mechanism of depression are criticized as either teleological, unspecific, or too abstract and elaborate to be correlated with clinical manifestations (chap. 17).

Beck's own theory of depression can be briefly summarized as follows: The individual becomes depression-prone during the developmental period, and acquires certain negative attitudes towards himself, the outside world and his future. These idiosyncratic cognitive patterns, when activated by others, constitute the primary and central factor of depression, with secondary disturbances of affect and behaviour. Depressive cognitive schemas displace more appropriate schemas, thought processes become illogical and are characterized by misinterpretation and distortion of reality, by perseveration and by loss of objectivity. Secondary affective responses increase the cognitive impairment yet further, leading to a vicious circle.

Professor Beck then makes a brief excursion to describe other psychopathological syndromes similarly in terms of their cognitive disorder.

In Part 5, after a survey of the results of psychopharmacology, ECT and psychotherapy, Professor Beck discusses cognitive (insight) therapy based on his own theory of depression.

The main feature of this treatment is to attempt to change the patient's cognitive organization during the symptom-free periods and thus to reduce his vulnerability to relapses.

The book is interesting and stimulating. Professor Beck's review of the literature is harsh, and he points out that much is merely 'alleged or taken for granted' without adequate evidence. Conveniently, the field is cleared for the presentation of his own work! Beck's theory of depression focuses on the problems of the mental function which is primarily affected in depression and he deduces all other disturbances from this. This theory evokes the traditions of 'mental faculties' and functional psychology. To the reviewers it would seem an arbitrary decision as to whether affect or cognition are primarily disturbed, since there appears to be no way of producing conclusive evidence for either alternative. But certainly it is important to pay attention to the role of the cognitive disorder in depression; and Professor Beck's analysis will be welcomed both by clinicians and by research workers.

TAMARA KOLAKOWSKA
PETER BEUMONT

Brain and Early Behaviour. Edited by R. J. Robinson. London: Academic Press, 1969. Pp. xvi + 374. £5.00.

This volume contains the Proceedings of the Second Study Group of the Centre for Advanced Study of Developmental Sciences. The original programme for the study group comprised only eight prepared papers. To these a further 13 contributions have been added resulting from (i) *ad hoc* papers given by members at the request of the group; (ii) discussions either of particular papers or of concepts and problems 'which the group felt the need to examine'. Following most of the contributions there is a short group discussion.

The book is divided into three main parts: The Foetus, The Infant, and Attempts at Synthesis. The part on the foetus consists of two papers, the first being a theoretical article by Bergstrom which attempts to link the anatomical ontogenesis of the brain, its electrical properties and its 'information transmitting' properties. His rather free use of terms from communication engineering and thermodynamics gives portentousness, but little conceptual illumination to his argument. The second contribution on the foetus is Tryphena Humphrey's hardy perennial on the postnatal recapitulation of prenatal reflex patterns.

The core of the book concerns the infant, or more correctly, the newborn, and some aspects of his development during the first year of life. The first section beings together summaries of work on nerve conduction velocity by Schulte, polygraphic analyses of behavioural states by Prechtl, and a detailed clinical description by Wolff, of an arrhinencephalic baby. The following sections deal with the development in infants of sleep (Petre-Quadens and Dittrichova), vision (Bower and Bronson) and learning (Lipsitt and Papousek).

Petre-Quadens is concerned with some parameters of paradoxical sleep in babies and their mothers before and after parturition. Similarities were found between these parameters in women at particular gestational stages and in premature babies of corresponding age. The author suggests that a common humoral mechanism may be responsible for the physiological similarities between mothers and children. The short papers by Dittrichova and Bronson were apparently off-the-cuff affairs. The former describes some changes in the sleep patterns of infants during the first 20 weeks of life. Bronson's paper reiterates the argument he advanced some years ago that the immature functional organization of the cerebral cortex in newborn infants does not support the claims made (e.g. by Fantz) for innate pattern perception. Interestingly, Bower, who follows Bronson, as well as adding a number of new tests to his original repertoire, is now more ready than one would have expected from his earlier writings to contemplate explanations of infant perception couched in terms of retinal variables.

The chapters by Prechtl, Lipsitt and Papousek have all appeared in very similar form in other symposia. In contrast, the single contribution to 'Primate Studies' by Ploog, which forms a kind of bridge passage to the final section (Synthesis), provides a pithy summary of work hitherto not published in its entirety in English.

The final three sections of the volume, 'Attempts at Synthesis', consist of discussions by individual speakers of what were thought to be key issues raised during the symposium, in particular, relationships between structure and function and the concept of 'organization' in behaviour. Not surprisingly, no very clear statement emerges on either issue. The matter is not helped by sticking doggedly to the procedure of following individual papers by group discussion, so that we get discussions of discussions! We are not therefore surprised to find that Wolff's 'summing-up'—described by the editor as an 'extempore tour de force'—is also followed by a group discussion so that the book weakly peters out discussing some minor issues of conditioning which could easily have been omitted or placed elsewhere. This is perhaps the only lapse in what, in my opinion, is the most distinguished feature of this volume: the great intelligence with which it has been edited. The relatively small proportion of prepared papers, combined with what still emerges as sheer prodigality in the choice of topics for discussion, must have presented enormous problems in terms of condensation, organization and structure. What Dr Robinson appears to have done is create papers out of long off-the-cuff statements, bring together discussion of particular topics under well-chosen subheadings and interject a short editorial of his own where necessary to anticipate terminological difficulties in the ensuing discussion. Few published symposia can owe so much to their editor.

It may be ungracious therefore to ask whether Dr Robinson's efforts were justified. However, much of the material is already published elsewhere, none of the papers are comprehensive reviews, what is new is of debatable importance, and the discussions re-state rather than offer solutions to the problems of which we lesser mortals are already cognizant.

It is desirable that conferences should be unite-disciplinary, but this is all the more reason for carefully delimiting the area of discussion beforehand. The real problem of this book is that the range of topics discussed is far too wide for coherence. The rubric 'Brain and Early Behaviour' sounds punchy, but it can be, and in this case has been, interpreted so widely that it becomes merely a rallying cry before each of the neurologists, physiologists, paediatricians, psychologists and psychiatrists, who form this group gallops off on his own particular hobby-horse. And one of the well-known facts of hobby-horse ethology is that they form stable social groups only under severe external pressure. Is it any wonder that some 15 per cent of the book is devoted (has to be devoted?) to 'Attempts at Synthesis'? How much better had the synthesis been made by the organizers before the conference was convened. In short, this book is a hotch-potch and no amount of skilful editing can disguise the fact.

S. J. HUTT

Behaviour Therapy. By Aubrey J. Yates. New York and London: Academic Press, 1970. Pp. 445. $13.95.

This is an impressive book. The literature on behavioural methods of treatment is growing rapidly. Yates has carried out the Herculean task of bringing this literature together, and in general he has done it well. Already, of course, the book is out of date, his review ends in 1968, but in such an expanding field this is inevitable.

This book is, however, more than a review of the literature, it is also a textbook of psychological treatment from the behaviourist's point of view, with important sections on theory and methodology. In the reviewer's opinion it is the best comprehensive text on this subject to date and will prove valuable to psychiatrist and clinical psychologist.

Having said that, there are various aspects which invite criticism. This is a book which was clearly intended to provoke discussion and it will probably do so with fruitful results.

Yates is at pains to point out, firstly, that behaviour therapy is distinguished from other forms of therapy in one way only, by its stress on the experimental investigation of the single case, and secondly that its essence is not theoretical but methodological.

The second point is an acceptable one; it is high time that the role of 'modern learning theory' was put into a proper perspective. The first point is a more doubtful one. It raises at least two questions. How important is it to distinguish behaviour therapy from other types of therapy, or to write a book on that subject alone, and secondly, are single case studies as important as Yates claims.

Relevant to the first question is Yates' section on the historical development of behaviour therapy and the role of the clinical psychologist. He discusses with admirable frankness the unsatisfactory way in which clinical psychologists are employed as 'pseudo-psychiatrists', or ancillary aids to diagnosis, insecure in their relationship with their dominating psychiatric colleagues, or 'basking in the reflected glory and status of the medical profession'. Perhaps as a reaction to this he concludes that the most important role of the clinical psychologist is that of 'fundamental research worker' in which his specialist knowledge of psychology is going to be best exploited. Undoubtedly there are many clinical psychologists who are of enormous value to their medical colleagues in designing and carrying out research. But there are many who are bad at research and many psychiatrists who are good at it and there is more to a good research worker than the type of basic training he has received. There are in addition many other things that a psychologist can do as well as or better than a psychiatrist and just because they will be done in a similar way does not mean that their training is being wasted. If, however, one connects Yates' concept of the clinical psychologist's role with his definition of behaviour therapy a possible explanation of his attitude emerges. Behaviour therapy is essentially an experimental method, the psychologist's essential role is the experimenter, hence behaviour therapy is primarily the province of the clinical psychologist. If one goes on to accept that this is an emotional rather than a rational viewpoint, the need to purify behaviour therapy and set it apart becomes easier to understand. The average psychiatrist is not concerned with the purity of behaviour therapy, he is interested in something that works and he will, in pragmatic fashion, incorporate useful principles from behaviour therapy into his general therapeutic approach. Perhaps Yates is concerned that the important contribution that psychologists have undoubtedly made in this area will become lost in the body of psychiatric knowledge and psychologists will lose the glory that should rightfully be theirs. Perhaps such an assessment does Yates an injustice but it fits the case and it enables several other inconsistent attitudes to be understood.

Yates emphasizes the importance of the individual case study and is very critical of the value of group comparisons. Homogeneity of groups, he says, is usually insufficient to permit valid comparison of different techniques and if it was sufficient it would paradoxically be impossible to generalize from the samples used. Somehow he persuades himself and possibly some of his readers that these problems are less if individual case studies are used; generalization is made to other patients similar to the one studied, he comments with bland optimism.

It is indeed possible that if, instead of treating a roughly homogeneous sample of patients with a standard technique individual studies were carried out, each one leading to the next, more would be learnt from the results. To do this in the way Yates implies would mean using each patient to test a hypothesis arising from the previous case and would require selection of patients and method so that only one variable changed each time. It is hardly surprising that no report of such a series of individual cases has so far appeared. The practical problems would be enormous and the time needed considerable. Patients with problems are not objects that can be kept in the drawer until they are needed. One can only conclude from the practical naivety of Yates' ideas that he has not attempted to put them into effect—a failing which has bedevilled too many writers in this field.

It would be incorrect to belittle the value of individual case studies, but they only answer some of the questions we need to pose and surely a happy mixture of single case and group studies is what is required. Very often a well-conducted single case will give rise to hypotheses that can be must efficiently tested in a group of patients. Unfortunately, to date there has been a large number of single case studies, some good, some bad, which have given rise to hypotheses which have not yet been put to the test and one possible consequence of this glamourizing of the single case study is that there will be a further burgeoning of premature publications which already clutter the literature with undigested and untried ideas. It is true that Yates is championing the 'proper' single case study, but all too often in this book he is prepared to attach importance to findings from studies which are improper by his own standards, and at times he seems to be biased against the group study which is well carried out whilst in favour of single case reports which are badly carried out.

Part of Yates' rejection of group comparison studies stems from his dislike of the medical model. Most physicians are well aware of the limitations of the medical model and it is true that it has become rooted in medical thinking. But its continued use is not merely habit, for there is a lack of any preferable alternative. Yates presumably has discovered the communication value of diagnostic categories as his whole book is planned according to the medical model; each chapter is a diagnostic category. He goes on, however, to conclude that comparing groups of patients is the same as comparing diagnostic categories. Obviously a group will be chosen on the basis of some common factor amongst its members, but that factor may be behavioural, diagnostic or what you will. Perhaps once again an emotive reaction to the medical profession underlies his attitude. It might be said that at such places as the Maudsley Hospital the psychiatrists have kept opportunities for group studies to themselves, leaving the psychologists to glean what they can from single cases that fall by the wayside. Even if that were so it would not be an indictment of group methods but of psychiatrists.

This review accuses Yates of an emotional bias which it is suggested influences not only his discussion of general principles but also his criticism of particular studies. Those that get 'good results' in general have more favourable reception and too often their theoretical premises are accepted far too readily. To be fair, however, this book must be compared with others that have gone before it. The history of behaviour therapy is a good example of the dialectic process with the clinical psychologist providing the antithesis. In this book Yates has come closer to a balanced synthesis than most of his psychological colleagues so far.

JOHN BANCROFT

CHANGE-OVER TO SI UNITS

From January 1972 the journals of the British Psychological Society will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Further information about SI Units is contained in the revised *Suggestions to Authors* pamphlet issued by the Society and obtainable, price 25p (U.S.A. $1.00) post free, from Cambridge University Press.

Authors who refer to physical measurements in their papers should now use SI Units except where it would be absurd to do so. Conversion tables will be published during the change-over period but in *Br. J. soc. clin. Psychol.* these will be restricted to units of length as others rarely occur. When they do, conversion factors will be stated in footnotes. Common units of time (e.g. hour, year) will persist.

1 metre (m) = 39.37 inches
1 metre (m) = 3.28 feet
0.0254m = 1 inch
0.3048m = 1 foot

REHABILITATION-INDIA
CALCUTTA
700063